

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG**

-----o0o-----



**ISO 9001 : 2008**

# **ĐỒ ÁN TỐT NGHIỆP**

## **NGÀNH CÔNG NGHỆ THÔNG TIN**

**HẢI PHÒNG 2017**

Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

---

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG**

-----o0o-----

**GHÉP ẢNH PANORAMA DỰA TRÊN ĐỐI  
SÁNH CÁC ĐẶC TRƯNG BẤT BIẾN**

**ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY**

**Ngành: Công nghệ Thông tin**

Sinh viên thực hiện: Lương Văn Kiên

Mã số sinh viên: 1312101024

Cán bộ hướng dẫn: Ts. Ngô Trường Giang

**HẢI PHÒNG – 2017**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**      **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  
**TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG**      *Độc lập - Tự do - Hạnh phúc*  
-----o0o-----

## **NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP**

Sinh viên: Lương Văn Kiên

Mã sinh viên: 1312101024

Lớp: CT1701

Ngành: Công nghệ Thông tin

Tên đề tài: Ghép ảnh Panorama dựa trên đối sánh các đặc trưng bất biến

## NHIỆM VỤ ĐỀ TÀI

1. Nội dung và các yêu cầu cần giải quyết trong nhiệm vụ đề tài tốt nghiệp

a. Nội dung

b. Các yêu cầu cần giải quyết

2. Các số liệu cần thiết để thiết kế, tính toán

3. Địa điểm thực tập

**CÁN BỘ HƯỚNG DẪN ĐỀ TÀI TỐT NGHIỆP**

**Người hướng dẫn thứ nhất:**

Họ và tên: Ngô Trường Giang

Học hàm, học vị: Tiến Sĩ

Cơ quan công tác: Trường Đại Học Dân Lập Hải Phòng

Nội dung hướng dẫn:

.....  
.....  
.....

**Người hướng dẫn thứ hai:**

Họ và tên:

Học hàm, học vị:

Cơ quan công tác:

Nội dung hướng dẫn:

.....  
.....  
.....

Đề tài tốt nghiệp được giao

ngày tháng năm 2017

Yêu cầu phải hoàn thành trước

ngày tháng năm 2017

Đã nhận nhiệm vụ: Đ.T.T.N

Đã nhận nhiệm vụ: Đ.T.T.N

**Sinh viên**

**Cán bộ hướng dẫn Đ.T.T.N**

**Ts. Ngô Trường Giang**

*Hải Phòng, ngày ..... tháng ..... năm 2017*

**HIỆU TRƯỞNG**

***GS.TS.NGUYỄN Trần Hữu Nghị***

**PHÂN NHẬN XÉT TÓM TẮT CỦA CÁN BỘ HƯỚNG DẪN**

**1. Tinh thần thái độ của sinh viên trong quá trình làm đề tài tốt nghiệp:**

.....

.....

.....

.....

.....

.....

**2. Đánh giá chất lượng của đề tài tốt nghiệp (so với nội dung yêu cầu đã đề ra trong nhiệm vụ đề tài tốt nghiệp)**

.....

.....

.....

.....

.....

.....

.....

**3. Cho điểm của cán bộ hướng dẫn:**

**(Điểm ghi bằng số và chữ)**

.....

.....

Ngày ..... tháng ..... năm 2017

**Cán bộ hướng dẫn chính**

*(Ký, ghi rõ họ tên)*

**PHÂN NHẬN XÉT ĐÁNH GIÁ CỦA CÁN BỘ CHẤM PHẢN BIỆN ĐỀ  
TÀI TỐT NGHIỆP**

**1. Đánh giá chất lượng đề tài tốt nghiệp (về các mặt như cơ sở lý luận, thuyết minh chương trình, giá trị thực tế, ...)**

.....  
.....  
.....  
.....  
.....  
.....

**2. Cho điểm của cán bộ phản biện  
(Điểm ghi bằng số và chữ)**

.....  
.....

*Ngày ..... tháng ..... năm 2017*

**Cán bộ chấm phản biện**

*(Ký, ghi rõ họ tên)*

## LỜI CẢM ƠN

Để hoàn thành đồ án này, ngoài sự cố gắng của bản thân, em xin chân thành cảm ơn các thầy giáo, cô giáo trong ban giám hiệu nhà trường, ban chủ nhiệm khoa cùng các thầy, cô giáo trong khoa Công nghệ thông tin – Trường đại học Dân Lập Hải Phòng đã tạo mọi điều kiện thuận lợi cho em trong quá trình học tập tại trường. Đặc biệt, em xin chân thành cảm ơn sự hướng dẫn tận tình của TS.Ngô Trường Giang – giảng viên khoa Công nghệ thông tin trường Đại học Dân Lập Hải Phòng đã tạo mọi điều kiện giúp đỡ em hoàn thành đồ án.

Em xin gửi lời cảm ơn chân thành đến các bạn lớp CT1701 (Khóa 2013-2017) đã đồng viên tinh thần và giúp đỡ em trong cuộc sống cũng như trong quá trình học tập.

Cuối cùng em xin gửi lời cảm ơn đặc biệt nhất tới gia đình, bố, mẹ, những người đồng viên, khích lệ để giúp em hoàn thành đồ án này.

Em rất mong nhận được những sự góp ý của thầy cô giáo và các bạn sinh viên để đề tài của em được hoàn thiện hơn.

**Em xin chân thành cảm ơn!**

Hải Phòng, ngày 27 tháng 12 năm 2017

*Người thực hiện*  
Lương Văn Kiên



**MỤC LỤC**

<b>LỜI CẢM ƠN.....</b>	<b>1</b>
<b>MỤC LỤC .....</b>	<b>9</b>
<b>DANH MỤC HÌNH ẢNH.....</b>	<b>11</b>
<b>MỞ ĐẦU.....</b>	<b>12</b>
<b>CHƯƠNG 1: TỔNG QUAN VỀ ĐỐI SÁNH ẢNH.....</b>	<b>14</b>
1.1 Tổng quan về ảnh số .....	14
1.1.1 Khái niệm về ảnh số.....	14
1.1.2 Điểm ảnh .....	14
1.1.3 Mức xám của ảnh .....	15
1.1.4 Lược đồ mức xám .....	15
1.1.5 Độ phân giải của ảnh.....	16
1.2 Một số vấn đề trong xử lý ảnh .....	16
1.2.1 Biến đổi ảnh .....	16
1.2.2 Biểu diễn ảnh.....	16
1.2.3 Phân tích ảnh.....	17
1.2.4 Nhận dạng ảnh.....	17
1.2.5 Nén ảnh .....	18
1.3 Các đặc trưng của ảnh số .....	18
1.3.1 Đặc trưng về màu sắc.....	19
1.3.2 Đặc trưng kết cấu .....	19
1.3.3 Đặc trưng hình dạng.....	19
1.3.4 Đặc trưng cục bộ bất biến .....	20
1.4 Đối sánh ảnh.....	21
1.4.1 Giới thiệu về đối sánh ảnh .....	21
1.4.2 Các phương pháp đối sánh ảnh .....	22
1.4.3 Đối sánh dựa theo đặc trưng .....	23
<b>CHƯƠNG 2: GHÉP ẢNH PANORAMA DỰA TRÊN ĐỐI SÁNH ĐẶC TRƯNG BẤT BIẾN .....</b>	<b>26</b>
2.1 Tổng quan về ghép ảnh.....	26

## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

---

2.1.1	Giới thiệu về ghép ảnh .....	26
2.1.2	Các kiểu ghép ảnh .....	27
2.1.3	Quá trình ghép ảnh Panorama .....	29
2.1.4	Các kỹ thuật ghép ảnh Panorama .....	36
2.2	Ghép ảnh Panorama dựa trên đặc trưng bất biến của ảnh .....	38
2.2.1	Trích chọn đặc trưng bất biến của ảnh .....	38
2.2.2	Đối sánh các đặc trưng bất biến .....	43
2.2.3	Tính toán ma trận Homography .....	45
2.2.4	Ghép ảnh dựa trên ma trận Homography .....	50
<b>CHƯƠNG 3:</b>	<b>THỰC NGHIỆM TẠO ẢNH PANORAMA .....</b>	<b>52</b>
3.1	Môi trường cài đặt.....	52
3.2	Giao diện chương trình .....	53
3.3	Chạy chương trình thực nghiệm .....	53
3.4	Kết quả chạy thực nghiệm .....	61
<b>KẾT LUẬN.....</b>	<b>.....</b>	<b>66</b>
<b>MỘT SỐ TÀI LIỆU THAM KHẢO .....</b>	<b>.....</b>	<b>67</b>

## DANH MỤC HÌNH ẢNH

- Hình 1.1: Ảnh đầu vào được thể hiện trên lược đồ xám  
Hình 2.1: Ví dụ về ảnh khảm  
Hình 2.2: Máy ảnh panorama  
Hình 2.3: Ví dụ về ảnh Panorama  
Hình 2.4: Máy ảnh được đặt trên một giá trượt  
Hình 2.5: Mô hình chụp ảnh có ván trượt  
Hình 2.6: Ví dụ cho recognize panorama  
Hình 2.7: Ảnh panorama chưa được trộn màu  
Hình 2.8: Ảnh panorama sau khi được trộn màu  
Hình 2.9: Ảnh panorama kết quả  
Hình 2.10: Ảnh panorama sau khi được cắt  
Hình 2.11: Ví dụ về kết cấu nhân tạo  
Hình 2.12: Ví dụ về kết cấu tự nhiên  
Hình 2.13: Cửa sổ trượt phát hiện góc Harris  
Hình 2.14: Minh họa các trường hợp  $\lambda_1$  và  $\lambda_2$   
Hình 2.15: Bộ mô tả cục bộ  
Hình 2.16: Ví dụ về đối sánh hai tập đặc trưng  
Hình 2.17: Phép chiếu Homography  
Hình 2.18: Minh họa ghép nối ảnh  
Hình 3.1: Giao diện chính của chương trình  
Hình 3.2: Ảnh đầu vào thứ nhất  
Hình 3.3: Ảnh đầu vào thứ hai  
Hình 3.4: Ảnh đầu vào thứ ba  
Hình 3.5: Hộp thoại chọn ảnh để ghép  
Hình 3.6: Hình ảnh được chọn sẽ hiển thị trên giao diện  
Hình 3.7: Click “Stitch image” để tiến hành ghép ảnh  
Hình 3.8: Kết quả tìm kiếm góc cho ảnh đầu vào thứ nhất  
Hình 3.9: Kết quả tìm kiếm góc cho ảnh đầu vào thứ hai  
Hình 3.10: Kết quả tìm kiếm góc cho ảnh đầu vào thứ ba  
Hình 3.11: Đối sánh ảnh thứ nhất và ảnh thứ hai  
Hình 3.12: Đối sánh ảnh thứ hai và ảnh thứ ba  
Hình 3.13: Ảnh thứ nhất bị biến đổi theo ảnh thứ hai  
Hình 3.14: Ảnh thứ hai làm tâm nên không bị biến đổi  
Hình 3.15: Ảnh thứ ba bị biến đổi theo ảnh thứ hai  
Hình 3.16: Ảnh panorama kết quả  
Hình 3.17: Hai ảnh có tỷ lệ trùng nội dung thấp  
Hình 3.20: Hai ảnh đầu vào có vị trí đứng chụp khác nhau  
Hình 3.20: Hai ảnh đầu vào có vị trí đứng chụp khác nhau  
Hình 3.22: Hai ảnh đầu vào có vị trí lệch nhau nhiều  
Hình 3.19: Hai ảnh kết quả với hai ngưỡng đối sánh khác nhau

## MỞ ĐẦU

Xử lý ảnh số có nhiều ứng dụng trong thực tế. Một trong những ứng dụng sớm nhất là xử lý ảnh từ nhiệm vụ Ranger 7 tại phòng thí nghiệm Jet Propulsion vào những năm đầu của thập kỷ 60. Hệ thống chụp hình gắn trên tàu vũ trụ có một số hạn chế về kích thước và trọng lượng, do đó ảnh nhận được bị giảm chất lượng như mờ, méo hình học và nhiễu nền. Các ảnh đó được xử lý thành công nhờ máy tính số. Hình ảnh của mặt trăng và sao hỏa mà chúng ta thấy trong các tạp chí đều được xử lý bằng máy tính số.

Bên cạnh ngôn ngữ giao tiếp, các thông tin dưới dạng hình ảnh đóng một vai trò rất quan trọng trong việc trao đổi thông tin. Trong công nghệ thông tin, xử lý ảnh và đồ họa đã chiếm một vị trí rất quan trọng bởi vì các đặc tính đầy hấp dẫn đã tạo nên một sự phân biệt với các lĩnh vực khác. Ta biết rằng phần lớn các thông tin mà con người thu thập được qua thị giác đều bắt nguồn từ các ảnh. Do đó việc xử lý ảnh và đồ họa là một bộ phận quan trọng trong việc trao đổi thông tin giữa người và máy.

Trong cuộc sống hiện đại ngày nay, người máy càng đóng vai trò quan trọng trong công nghiệp và gia đình. Chúng sẽ thực hiện những công việc rất nhàm chán hoặc nguy hiểm, và những công việc mà tốc độ và độ chính xác vượt quá khả năng của con người. Khi người máy trở nên tinh vi hơn, thị giác máy tính sẽ đóng vai trò ngày càng quan trọng. Người ta sẽ đòi hỏi người máy không những phát hiện và nhận dạng các bộ phận công nghiệp, mà còn hiểu được những gì chúng thấy và đưa ra những hành động phù hợp. Xử lý ảnh sẽ tác động lớn đến thị giác máy tính.

Những ứng dụng khác của xử lý ảnh là vô cùng đa dạng. Ngoài những ứng dụng đã thảo luận ở trên, còn bao gồm cả các lĩnh vực khác như điện tử gia đình, thiên văn học, sinh vật học, vật lý, nông nghiệp, nhân chủng học, ...

Đặc biệt, xử lý ảnh còn được ứng dụng trong ghép ảnh để tạo ra những bức ảnh có chiều rộng và chiều sâu mà khi chụp máy ảnh không cho phép góc nhìn rộng như thế.

Ngày nay, hầu hết các loại máy ảnh thông thường dù độ phân giải cao nhưng cũng chỉ có thể ghi lại được một phần của những đối tượng có kích thước lớn như công viên hay một thành phố. Do vậy yêu cầu được đặt ra là phải làm như thế nào để có thể ghép được các tấm ảnh nhỏ đó thành một tấm ảnh lớn hiển thị đầy đủ các đối tượng có kích thước lớn đó. Đây cũng chính là lý do mà em chọn chủ đề ghép ảnh panorama dựa trên đối sánh đặc trưng trong đề án.

## CHƯƠNG 1: TỔNG QUAN VỀ ĐỐI SÁNH ẢNH

### 1.1 Tổng quan về ảnh số

#### 1.1.1 Khái niệm về ảnh số

Ảnh số là tập hợp hữu hạn các điểm ảnh với mức xám phù hợp dùng để mô tả gần nhất với ảnh thật. Số điểm ảnh xác định độ phân giải của ảnh, độ phân giải càng cao thì càng thể hiện rõ nét các đặc điểm của tấm hình, càng làm cho tấm ảnh trở nên thực và sắc nét hơn.

Ảnh có thể được biểu diễn theo một trong hai mô hình: mô hình Vector hoặc mô hình Raster.

- Mô hình Vector: Ngoài mục đích tiết kiệm không gian lưu trữ, dễ dàng hiển thị và in ấn, các ảnh biểu diễn theo mô hình vector còn có ưu điểm cho phép dễ dàng lựa chọn, sao chép, di chuyển, tìm kiếm... Trong mô hình này, hướng vector của các điểm ảnh lân cận được sử dụng để mã hóa và tái tạo lại hình ảnh ban đầu. Các ảnh vector được thu nhận trực tiếp từ các thiết bị số hóa như Digitalize hoặc được chuyển đổi từ các ảnh Raster thông qua các chương trình vector hóa.
- Mô hình Raster: là mô hình biểu diễn ảnh thông dụng nhất hiện nay. Ảnh được biểu diễn dưới dạng ma trận các điểm ảnh. Tùy theo nhu cầu thực tế mà mỗi điểm ảnh có thể được biểu diễn bởi một hay nhiều bit. Mô hình Raster thuận lợi cho việc thu nhận, hiển thị và in ấn. Các ảnh được sử dụng trong phạm vi của đề tài này cũng là các ảnh được biểu diễn theo mô hình Raster.

#### 1.1.2 Điểm ảnh

Điểm ảnh (Pixel) là một phần tử của ảnh số tại tọa độ  $(x, y)$  với độ xám hoặc màu nhất định. Kích thước và khoảng cách giữa các điểm ảnh đó được chọn thích hợp sao cho mắt người cảm nhận được sự liên tục về không gian và mức xám (hoặc màu) của ảnh số gần như ảnh thật. Mỗi phần tử trong ma trận được gọi là một phần tử ảnh.

### 1.1.3 Mức xám của ảnh

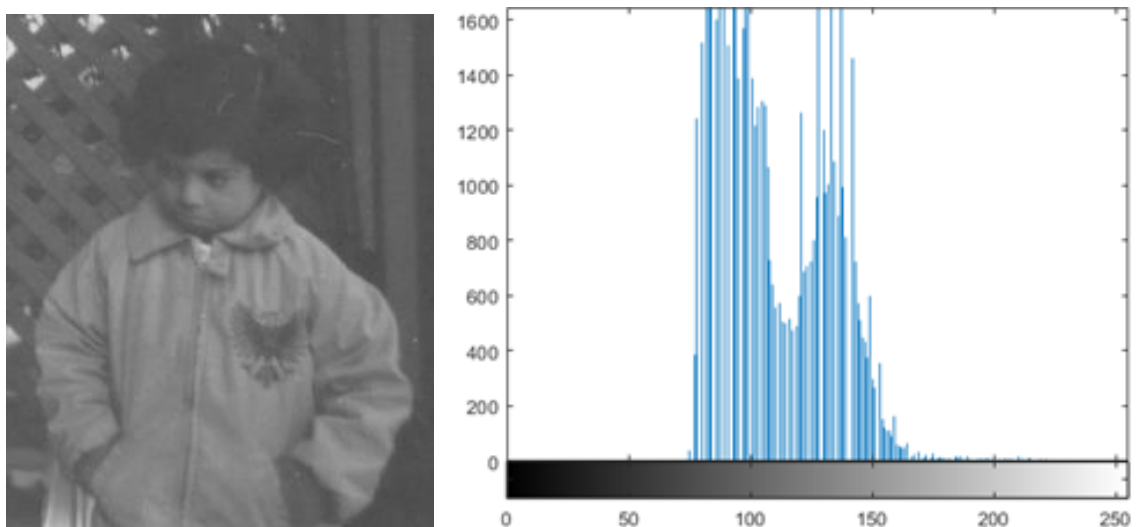
Là kết quả của sự biến đổi tương ứng một giá trị độ sáng của một điểm ảnh với một giá trị nguyên dương. Thông thường nó xác định trong khoảng từ 0 đến 255 tùy thuộc vào giá trị mà mỗi điểm ảnh được biểu diễn.

### 1.1.4 Lược đồ mức xám

Lược đồ mức xám (Histogram) hay còn gọi là lược đồ xám của một ảnh là một hàm cung cấp tần suất xuất hiện của mỗi mức xám (grey level).

Lược đồ xám của một ảnh có các mức xám trong khoảng  $[0, L-1]$  là một hàm rời rạc  $p(r_k) = n_k/n$ . Trong đó  $n_k$  là số pixel mức xám thứ  $r_k$ ,  $n$  là tổng số pixel của ảnh và  $k = 1, 2, 3, \dots, L-1$ . Vẽ hàm này với tất cả các giá trị của  $k$  sẽ biểu diễn khái quát sự xuất hiện các mức xám của một ảnh. Lược đồ mức xám của ảnh có thể được biểu diễn thông qua tần suất xuất hiện mỗi mức xám trên hệ tọa độ vuông góc Oxy. Trong đó, trục hoành biểu diễn số mức xám từ 0 đến  $N$  (số bit của ảnh xám), trục tung biểu diễn số pixel của mỗi mức xám.

Nhìn vào biểu đồ có thể biết được phân bố cường độ sáng của một ảnh, với những ảnh mà phân bố histogram lệch về bên phải thì ảnh đó là một ảnh có độ sáng tốt, ngược lại thì ảnh đó là một ảnh tối.



**Hình 1.1: Ảnh đầu vào được thể hiện trên lược đồ xám**

### 1.1.5 Độ phân giải của ảnh

Độ phân giải (Resolution) của ảnh là mật độ điểm ảnh được ấn định trên một ảnh số được hiển thị. Như trình bày ở trên, khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân bố, đó chính là độ phân giải và được phân bố theo trục x và y trong không gian hai chiều.

## 1.2 Một số vấn đề trong xử lý ảnh

### 1.2.1 Biến đổi ảnh

Thuật ngữ biến đổi ảnh thường được dùng để nói tới một lớp các ma trận đơn vị và các kỹ thuật dùng để biến đổi ảnh. Cũng như các tín hiệu một chiều được biểu diễn bởi một chuỗi các hàm cơ sở, ảnh cũng có thể được biểu diễn dưới một số chuỗi rời rạc các ma trận cơ sở gọi là ảnh cơ sở. Phương trình ảnh cơ sở có dạng:

$$A^*_{k,l} = a_k a_l^{*T}$$

Với  $a_k$  là cột thứ k của ma trận A. A là ma trận đơn vị. Có nghĩa là  $AA^{*T}=1$ . Các  $A^*_{k,l}$  được định nghĩa ở trên với  $k, l = 0, 1, 2, \dots, N-1$  là ảnh cơ sở. Có nhiều loại biến đổi được dùng như:

- Biến đổi Fourier, Sin, Cosin, Hadamard....
- Tích Kronecker.
- Biến đổi KL (Krhumen loeve).

Do phải xử lý nhiều thông tin, các phép toán nhân và cộng trong khai triển là quá lớn, nên các phép biến đổi trên nhằm giảm thứ nguyên của ảnh để việc xử lý ảnh được hiệu quả hơn.

### 1.2.2 Biểu diễn ảnh

Trong biểu diễn ảnh, người ta thường dùng các phần tử đặc trưng của ảnh là pixel. Các mô hình biểu diễn ảnh cho thấy một mô tả logic hay định



lượng các tính chất của hàm này. Trong biểu diễn ảnh cần chú ý đến tính trung thực của ảnh hoặc các tiêu chuẩn để đo chất lượng ảnh hoặc tính hiệu quả của các kỹ thuật xử lý.

Việc xử lý ảnh số yêu cầu ảnh phải được mẫu hóa và lượng tử hóa. Việc lượng tử hóa ảnh là chuyển đổi tín hiệu tương tự sang tín hiệu số của một ảnh đã lấy mẫu sang một số hữu hạn mức xám.

Một số mô hình thường được dùng trong biểu diễn ảnh: Mô hình toán, mô hình thống kê. Trong mô hình toán, ảnh hai chiều được biểu diễn nhờ các hàm hai biến trực giao gọi là các hàm cơ sở. Với mô hình thống kê, một ảnh được coi như một phần tử của một tập hợp đặc trưng bởi các đại lượng như: kỹ vọng toán học, hiệp biến, phương sai, moment.

### 1.2.3 Phân tích ảnh

Phân tích ảnh liên quan đến việc xác định các độ đo định lượng của một ảnh để đưa ra một mô tả đầy đủ về ảnh. Các kỹ thuật được sử dụng nhiều nhất là các kỹ thuật phát hiện biên của ảnh, ví dụ như lọc vi phân hay dò theo quy hoạch động. Bên cạnh đó, người ta cũng có thể dùng các kỹ thuật để phân vùng ảnh. Từ ảnh thu được tiến hành kỹ thuật tách (split) hay hợp (fusion) dựa theo các tiêu chuẩn đánh giá như: màu sắc, cường độ, vv... Các phương pháp được biết đến như Quad-Tree, mảnh hóa biên, nhị phân hóa đường biên. Cuối cùng, phải kể đến các kỹ thuật phân lớp dựa theo cấu trúc.

### 1.2.4 Nhận dạng ảnh

Nhận dạng ảnh là quá trình liên quan đến các mô tả đối tượng mà người ta muốn đặc tả nó. Quá trình nhận dạng thường đi sau quá trình trích chọn các đặc tính chủ yếu của đối tượng:

- Mô tả tham số (nhận dạng theo tham số).
- Mô tả theo cấu trúc (nhận dạng theo cấu trúc).

Trên thực tế, con người đã áp dụng kỹ thuật nhận dạng khá thành công với nhiều đối tượng khác nhau như: nhận dạng ảnh vân tay, nhận dạng chữ (chữ cái, chữ số, chữ có dấu).

### 1.2.5 Nén ảnh

Dữ liệu ảnh cũng như các dữ liệu khác cần phải lưu trữ hay truyền đi trên mạng. Như đã nói ở trên, lượng thông tin để biểu diễn cho một ảnh là rất lớn. Do đó làm giảm lượng thông tin hay nén dữ liệu là một nhu cầu cần thiết.

Phân loại phương pháp nén bao gồm:

- Dựa vào nguyên lý nén:
  - Nén không mất thông tin: Sau khi giải nén ta thu được chính xác dữ liệu gốc.
  - Nén có mất thông tin: Sau khi nén ta không thu được dữ liệu như bản gốc.
- Dựa vào cách thức thực hiện nén:
  - Phương pháp không gian: Tác động trực tiếp lên việc lấy mẫu của ảnh trong miền không gian.
  - Phương pháp sử dụng biến đổi: Tác động lên sự biến đổi của ảnh gốc.
- Dựa vào triết lý của sự mã hóa:
  - Các phương pháp nén thể hệ thứ nhất: Gồm các phương pháp mà mức độ tính toán là đơn giản.
  - Các phương pháp nén thể hệ thứ hai: Dựa vào độ bão hòa của tỷ lệ nén.

### 1.3 Các đặc trưng của ảnh số

Trong phạm vi xử lý ảnh, đặc trưng ảnh số là một phần thông tin ảnh số thích hợp cho các nhiệm vụ tính toán liên quan đến một ứng dụng nhất định.

Những đặc trưng đó có thể là kết cấu đặc biệt trong ảnh số như các điểm, các cạnh của một đối tượng hoặc một đối tượng nào đó trong ảnh. Mặt khác, các đặc trưng của ảnh số cũng có thể là kết quả của một phép biến đổi toàn diện hoặc là các phương pháp phát hiện điểm đặc trưng được áp dụng trên toàn bộ ảnh đó.

Điểm đặc trưng trong ảnh là một điểm ảnh có chứa nhiều thông tin hơn các điểm ảnh lân cận. Biểu diễn ảnh theo điểm đặc trưng sẽ cô đọng hơn, giảm được không gian tìm kiếm trong các bài toán ứng dụng.

### 1.3.1 Đặc trưng về màu sắc

Là một đặc trưng nổi bật và được sử dụng phổ biến nhất trong các ứng dụng xử lý ảnh [3]. Mỗi một điểm ảnh (thông tin màu sắc) có thể biểu diễn trong không gian màu sắc 3 chiều. Các không gian màu sắc thường dùng là: RGB, CIE, HSV ...

Hiện nay các công cụ tìm kiếm như google, yahoo, bing ... đều dựa theo đặc trưng về màu sắc để tìm kiếm ảnh liên quan kết hợp với đặc trưng kết cấu và đặc trưng hình dạng.

### 1.3.2 Đặc trưng kết cấu

Kết cấu cung cấp thông tin về sự sắp xếp về mặt không gian của màu sắc và cường độ một ảnh. Kết cấu được đặc trưng bởi sự phân bố không gian của những mức cường độ trong một khu vực lân cận với nhau. Kết cấu gồm các kết cấu góc hay nhiều kết cấu gộp lại đôi khi gọi là texel.

Đặc trưng kết cấu được sử dụng rộng rãi và rất trực quan nhưng không có định nghĩa chính xác bởi tính biến thiên rộng của nó. Có rất nhiều cách để mô tả kết cấu: Những phương pháp thống kê thường sử dụng tần số không gian, ma trận biến cố, tần số biên, ...

### 1.3.3 Đặc trưng hình dạng

Hình dạng của một ảnh hay một vùng là một đặc trưng quan trọng trong việc xác định và phân biệt ảnh trong nhận dạng mẫu. Định nghĩa hình dạng

của đối tượng thường là rất khó. Hình dạng thường được biểu diễn bằng lời nói hoặc hình vẽ, và mọi người thường sử dụng thuật ngữ như là tròn, méo. Xử lý hình dạng dựa trên máy tính đòi hỏi rất phức tạp, trong khi rất nhiều phương pháp mô tả hình dạng thực tế đang tồn tại nhưng không có một phương pháp chung nào cho mô tả hình dạng. Có hai kiểu đặc trưng hình dạng chính thường được sử dụng:

- Những đặc trưng dựa trên biên: chỉ sử dụng đường bao ngoài của hình dạng.
- Những đặc trưng vùng: sử dụng toàn bộ vùng của hình dạng.

Mục tiêu chính của biểu diễn hình dạng trong nhận dạng mẫu là đo thuộc tính hình học của một đối tượng được dùng trong phân lớp, so sánh và nhận dạng đối tượng.

Độ đo về hình dạng có rất nhiều trong phạm vi lý thuyết xử lý ảnh. Chúng trải rộng từ những độ đo toàn cục dạng thô sơ trợ giúp cho việc nhận dạng đối tượng, cho tới những độ đo chi tiết tự động tìm kiếm những hình dạng đặc biệt.

### 1.3.4 Đặc trưng cục bộ bất biến

Là những điểm đặc trưng không thay đổi khi xoay ảnh, co giãn ảnh hay thay đổi cường độ sáng của ảnh. SIFT là đặc trưng bất biến được sử dụng rộng rãi:

- SIFT: Là viết tắt của cụm từ Scale-Invariant Feature Transform, là một trong những thuật toán nổi tiếng nhất hiện nay dùng để phát hiện và mô tả các đặc trưng ảnh số. Thuật toán này được công bố bởi David Lowe vào năm 1999.
- SURF: Là viết tắt của cụm từ Speeded Up Robust Features, được giới thiệu vào năm 2006 bởi nhóm các nhà nghiên cứu bao gồm Herbert Bay, Tinne Tuytelaars và Luc Van Gool. Được phát triển dựa trên thuật

toán SIFT nhưng được cải tiến để cho tốc độ xử lý nhanh hơn giải thuật SIFT.

Ở thuật toán SIFT, việc tìm scale-space dựa trên việc tính gần đúng LoG(Laplace of Gaussian) dùng DoG (Difference of Gaussian), trong khi đó SURF sử dụng Box Filter, tốc độ xử lý sẽ được cải thiện đáng kể với việc dùng ảnh tích phân (integral image). Ở bước xác định hướng, SURF sử dụng wavelet response theo hai chiều dọc và ngang, sau đó tìm hướng chính bằng cách tính tổng các response đó.

## **1.4 Đối sánh ảnh**

### **1.4.1 Giới thiệu về đối sánh ảnh**

Đối sánh ảnh là một bài toán đã và đang thu hút được sự quan tâm của các nhà nghiên cứu và phát triển[1]. Mỗi khi bài toán này được giải quyết, nó mở ra rất nhiều các ứng dụng hữu ích như: tìm kiếm ảnh, nhận dạng, theo dõi và phát hiện đối tượng, ghép ảnh, vv. Đối sánh hai ảnh là tìm ra những vùng giống nhau trên hai ảnh.

Thông thường, để đối sánh ảnh cần so sánh các phần tử cơ bản cấu thành nên nó. Đơn giản nhất là so sánh các điểm ảnh (pixel). Tuy nhiên phép so sánh này đòi hỏi nhiều thời gian tính toán và thường không đạt được độ chính xác như mong muốn.

Giải pháp đầu tiên cho vấn đề đối sánh ảnh được đề xuất bởi Hobrough vào cuối những năm 1950. Hệ thống tự động tìm kiếm các điểm liên hợp đầu tiên được giới thiệu bởi công ty Wild Heerbrugg năm 1964 nhưng lại không được sử dụng phổ biến. Tuy nhiên, ý tưởng của Hobrough áp dụng mối tương quan chéo lại được nhiều người sử dụng. Từ những năm 1970, việc tập trung phát triển đối sánh ảnh và đối sánh tương quan gặt hái được nhiều thành công và được áp dụng trong hệ thống đo độ tương tự cho ảnh (Helava, 1978). Ngày nay, công nghệ đối sánh ảnh được tích hợp trong nhiều phần mềm xử lý ảnh được sử dụng như là một công cụ tính toán. Có rất nhiều nghiên cứu được

thực hiện với mong muốn tìm hai điểm tương đồng trên hai bức ảnh. Thuật toán tìm kiếm điểm tương đồng có thể thực hiện được trên ảnh 2D.

Vấn đề chính của việc đối sánh ảnh là việc chọn một đối tượng phù hợp và cách thức để so sánh nó. So sánh theo từng pixel sẽ không khả thi với những ảnh có kích thước lớn vì nó sẽ cần tính toán nhiều hơn, mất nhiều thời gian hơn, hoặc muốn rút ngắn thời gian thì cần có phần cứng xử lý mạnh hơn. Hơn nữa dẫn đến sự không chính xác vì sự lặp đi lặp lại của các màu có cùng giá trị mức xám và nhiễu của ảnh. Để giải quyết vấn đề đó, thay vì đối sánh từng pixel một dẫn đến dữ liệu đầu vào quá lớn thì ta sẽ giảm dữ liệu đầu vào bằng cách đưa vào các đặc trưng của cả hai ảnh rồi tiến hành đối sánh trên các đặc trưng đó.

### 1.4.2 Các phương pháp đối sánh ảnh

#### 1.4.2.1 Đối sánh dựa theo vùng

Phương pháp này còn được gọi là phương pháp tương quan hay đối sánh mẫu. Phương pháp này kết hợp giữa đối sánh đặc trưng và đối sánh thành phần. Cường độ xám của ảnh được sử dụng làm cơ sở cho việc đối sánh ảnh. Do việc đối sánh từng pixel của cả hai ảnh là một việc bất khả thi nên thay vào đó, ta sẽ đối sánh một tập các điểm ảnh lân cận nhau để giảm số lần tính toán. Tại ảnh thứ nhất sử dụng một cửa sổ có kích thước  $m \times n$  (thông thường là  $m=n$  để có thể dễ dàng tìm được tọa độ điểm trung tâm của cửa sổ) đem so sánh với một “mẫu” cũng là cửa sổ có kích thước tương tự ở ảnh thứ hai. Các phép so sánh được thực hiện trên cửa sổ. Trong phép đo ảnh thì tương quan chéo và đối sánh bình phương tối thiểu là những kỹ thuật được sử dụng nhiều trong đối sánh ảnh dựa theo vùng.

Kích thước mẫu càng lớn thì yêu cầu về tính đặc trưng của thực thể được đối sánh càng cao. Mặt khác sự biến dạng hình học gây ra bởi hiện tượng xoay ảnh cũng sẽ ảnh hưởng tới kết quả đối sánh của các mẫu có kích thước lớn. Yêu cầu về tính đặc trưng của thực thể cũng không được thỏa mãn

nếu vùng đó bị lặp đi lặp lại hoặc độ tương phản và cấu trúc thấp (Ví dụ: cát sa mạc, nước biển). Những vùng bị che khuất bởi các đối tượng khác cao hơn cũng nên bị loại bỏ. Để có được kết quả chấp nhận được, kích thước của mẫu phải nhỏ hoặc hình dạng phải thích nghi với biến dạng hình học.

Để tránh sai lệch về kết quả đối sánh, vị trí của cửa sổ tìm kiếm phải được xác định chính xác trong đối sánh dựa theo vùng. Kích thước của cửa sổ tìm kiếm phụ thuộc vào vị trí chính xác và về độ biến dạng do hướng của ảnh.

Sau khi tìm ra vị trí phù hợp nhất thì cần đánh giá độ chính xác và độ tin cậy của kết quả đối sánh tìm được. Thiết lập ngưỡng cho các phép đối sánh là một biện pháp để giảm thiểu việc đối sánh bị sai lệch. Ngoài biện pháp sử dụng ngưỡng thì có thể sử dụng phương pháp điều chỉnh hình học để tính toán và loại trừ những kết quả đối sánh sai.

### 1.4.3 Đối sánh dựa theo đặc trưng

Trái ngược với phương pháp đối sánh dựa theo vùng phương pháp đối sánh dựa trên đặc trưng sử dụng sự biến đổi đột ngột về các giá trị mức xám tương ứng với các đặc trưng của ảnh làm cơ sở để đối sánh như cạnh, góc, hoặc điểm đặc trưng của ảnh. Kỹ thuật đối sánh dựa theo đặc trưng vượt trội hơn so với kỹ thuật đối sánh dựa theo vùng. Kỹ thuật đối sánh dựa vào đặc trưng của ảnh về cơ bản gồm 3 bước chính:

- Chọn các điểm là điểm đặc trưng của ảnh (cạnh, góc, điểm) trong mỗi ảnh độc lập.
- Xây dựng danh sách các cặp điểm có thể là tương đồng.
- Tiến hành đối sánh và trả về kết quả tập các điểm tương.

Thông thường người ta sẽ tích hợp cả đối sánh vùng và đối sánh điểm đặc trưng vào các phần mềm ghép ảnh để đạt được kết quả chính xác nhất và tốc độ xử lý nhanh hơn, tốn ít thời gian hơn. Với sự phát triển của công nghệ như hiện tại thì thực hiện đối sánh trên những ảnh cỡ nhỏ thì thời gian thực

hiện không đáng kể, nhưng đối với ảnh có kích thước lớn thì việc tối ưu thuật toán, cải thiện tốc độ thực hiện thuật toán cũng là một vấn đề cần được quan tâm.

#### 1.4.3.1 Điểm quan tâm (Interest points)

Đối sánh dựa trên đặc trưng của ảnh áp dụng tốt nhất trên các khu vực ảnh có độ tương phản cao. Những điểm có thể được mô tả bằng sự chênh lệch cao về giá trị mức xám hoặc có gradient dốc được gọi là điểm quan tâm. Các điểm quan tâm nên có sự khác biệt, bất biến đối với sự biến dạng hình học và chất lượng của bức ảnh và có tính ổn định. Việc tìm kiếm điểm quan tâm trong ảnh được thực hiện qua hai bước:

- Tính toán các đặc trưng ở mỗi cửa sổ của ảnh được chọn.
- So sánh giá trị vừa tìm được với một ngưỡng cho trước.

Đặc trưng khác nhau với mỗi toán tử khác nhau, nhưng về cơ bản đều dựa trên giá trị mức xám bên trong mỗi cửa sổ trượt. Chỉ có những cửa sổ mà có giá trị lớn hơn hoặc nhỏ hơn ngưỡng mới được chấp nhận là điểm quan tâm. Một danh sách các điểm quan tâm của mỗi ảnh được đối sánh với tọa độ điểm ảnh của nó (điểm trung tâm của mỗi cửa sổ trượt) và mô tả của chúng là kết quả của quá trình xử lý.

#### 1.4.3.2 Cạnh và vùng

Cạnh có thể mô tả như là việc thay đổi đột ngột giá trị mức xám trong một vùng nhỏ. Cạnh thường tương ứng với biên của đối tượng trong ảnh. Quá trình trích xuất cạnh rất phức tạp và trải qua 3 bước bao gồm:

- Xác định các điểm ảnh nằm trên cạnh, giá trị mức xám bị ngắt quãng sẽ được xác định bằng trung bình cộng của các toán tử cạnh. Điểm đó có được xác định là điểm nằm trên cạnh hay không dựa vào kết quả so sánh giá trị mức xám với một ngưỡng cho trước.
- Nối các điểm ảnh với nhau và làm liền biên.



- Nhóm các cạnh với nhau, phân đoạn.

Toán tử cạnh sẽ phát hiện ra sự thay đổi của giá trị mức xám trong ảnh, dựa trên phép đạo hàm bậc nhất để tìm ra cực trị và định vị điểm cạnh. Một số toán tử cạnh có thể dùng như toán tử Robert (Robert Cross), toán tử Sobel (Sobel Operator), toán tử Prewitt (Prewitt Operator). Toán tử Sobel sẽ ít bị ảnh hưởng bởi nhiễu của ảnh vì bao gồm cả những điểm ảnh lân cận.

Toán tử Laplacian dựa trên phép đạo hàm bậc hai. Để không bị ảnh hưởng bởi nhiễu thì nó được kết hợp với toán tử Gaussian để làm mịn ảnh, khử nhiễu.

## **CHƯƠNG 2: GHÉP ẢNH PANORAMA DỰA TRÊN ĐỐI SÁNH ĐẶC TRƯNG BẤT BIẾN**

### **2.1 Tổng quan về ghép ảnh**

#### **2.1.1 Giới thiệu về ghép ảnh**

Ghép ảnh là quá trình kết hợp nhiều hình ảnh nhỏ xếp chồng lên nhau để tạo ra một bức ảnh lớn có độ phân giải cao hơn. Thông thường việc ghép ảnh được thực hiện bằng việc sử dụng các phần mềm máy tính.

Ghép ảnh có rất nhiều ứng dụng khác nhau. Ứng dụng truyền thống nhất là tạo nên ảnh không gian rộng và ảnh vệ tinh từ một tập các ảnh, dùng để xây dựng bản đồ địa lý, ghép các tấm ảnh chụp được trên bề mặt của một ngôi sao thành một tấm ảnh có độ phân giải lớn hơn, vv...

Các vấn đề chính trong ghép ảnh là sắp xếp các ảnh thành phần, nắn chỉnh biến dạng, biến đổi màu sắc và làm mờ đường biên giữa các ảnh. Tất cả các thao tác này nhằm làm cho bức ảnh ghép trông giống như là một ảnh liền chứ không phải là được ghép từ nhiều ảnh nhỏ.

Một số phần mềm ghép ảnh đã và đang được phát triển: Easy Mosaic, AndreaMosaic, Mazaika, Autopanopro, Panorama Factory, vv....

Việc ghép các thành phần của các đối tượng lại với nhau để thu được các ảnh tương ứng hoàn thiện hơn là một công việc khó khăn rất nhiều khi phải làm thủ công, mặt khác các ảnh khi thu nhận để ghép thường hay bị lệch và biến dạng đi một khoảng nào đấy. Yêu cầu đặt ra cần xác định độ sai lệch về thông tin giữa các phần ảnh định ghép, sau đó hiệu chỉnh độ sai lệch và cuối cùng là ghép chúng lại. Nghiên cứu kỹ thuật ghép ảnh còn mở ra cho chúng ta một hướng phát triển mới trong tương lai đó là kỹ thuật giả lập 3D.

## 2.1.2 Các kiểu ghép ảnh

### 2.1.2.1 Khảm ảnh

Khảm ảnh là việc tạo ra hình ảnh mới bằng cách ghép các ảnh nhỏ vào một ảnh lớn sao cho khi nhìn tổng thể vào ảnh lớn thì ta vẫn có thể nhìn thấy nội dung của bức ảnh lớn trước đó.



**Hình 2.1: Ví dụ về ảnh khảm**

Từ bức tranh tổng thể ban đầu, bằng các kỹ thuật xử lý khác nhau thì các bức ảnh nhỏ được lồng ghép vào đó tạo nên bức ảnh mới. Tất nhiên là nếu nhìn một cách tổng thể thì nó vẫn chính là bức tranh lớn ban đầu có điều nó khác đi một chút bởi những chi tiết bên trong đã được thay thế bởi các hình ảnh đơn lẻ.

### 2.1.2.2 Ảnh toàn cảnh (panorama)

Ảnh toàn cảnh là một cách nhìn rộng của một vật chất trong không gian. Nó cho phép biểu thị một góc nhìn rộng của các bức tranh, bản vẽ đồ họa, nghệ thuật nhiếp ảnh, phim hoặc video, hay mô hình 3 chiều.

Thuật ngữ panorama xuất hiện trước khi chúng ta có máy ảnh panorama. Nguyên gốc của từ panorama được xác định là do họa sĩ người Ireland – Robert Baker dùng để mô tả những bức tranh diện rộng ở Edinburgh (Đức). Những bức tranh panorama này được cuộn trong một ống hình trụ và người ta kéo ra từ từ để bức tranh dần được hiển thị.

## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

---

Năm 1881, họa sĩ người Hà Lan – Hendrik Willem Mesdag đã tạo nên trường phái Panorama Mesdag với những ống hình trụ cuộn các bức tranh toàn cảnh với kích cỡ khổng lồ, cao 14m và dài có thể từ 40 – 120m. Thế kỷ XIX có hai bức tranh panorama được coi là lớn nhất thời kỳ này, đó là tranh mô tả trận chiến tại Atlanta với chiều cao gần 13m, dài 110m. Bức tranh được xác định lớn nhất là ở Wroclaw (Ba Lan) với kích thước là 15m x 120m.

Do nhu cầu của con người và sự phát triển của khoa học công nghệ, người ta đã sáng tạo ra máy ảnh panorama. Nếu máy ảnh thường chỉ có thể chụp ảnh với một góc 90 độ thì máy ảnh panorama có thể chụp với một góc 175 độ, 180 độ hoặc 360 độ. Trước một không gian rộng lớn, máy ảnh thường bất lực trong việc ghi lại hình ảnh ở một góc rộng nhưng máy panorama lại phát huy được tác dụng của nó. Máy panorama thường được chụp bằng phim dương bản (còn được gọi là film slide). Chụp xong có thể xem film là biết được ảnh sẽ được in ra như thế nào.

Chính vì góc ảnh của panorama rộng nên máy ảnh panorama không có ống kính dài như máy thường. Ống kính của máy panorama có hình vòng cung. Khi chụp ống kính sẽ quét từ trái sang phải nên chúng ta phải sử dụng chân máy khi chụp.



**Hình 2.2: Máy ảnh panorama**

Ảnh Panorama nói đơn giản là xem hình ảnh với góc nhìn rộng hơn ảnh bình thường, tức là những khung ảnh cực kỳ rộng lớn mà một khung hình chụp bằng máy ảnh không thể hiện hết. Ảnh được ghép từ những ảnh số chụp từng phần của một phong cảnh (những cảnh này có phần gối nhau) lại thành một ảnh toàn cảnh hoàn chỉnh.

Ta có thể hiểu một cách đơn giản panorama là chế độ chụp ảnh khổ rộng bằng cách chụp nhiều tấm ảnh liên tiếp, với thông tin tấm ảnh trước được thể hiện một phần trong tấm ảnh sau, nhằm hỗ trợ người dùng. Sau đó với sự trợ giúp của phần mềm xử lý ảnh thì ta sẽ có được một tấm ảnh khổ rộng.



**Hình 2.3: Ví dụ về ảnh Panorama**

### 2.1.3 Quá trình ghép ảnh Panorama

Quá trình ghép ảnh panorama trải qua bốn bước cơ bản [2]:

#### 2.1.3.1 Thu nhận ảnh

Giai đoạn đầu tiên của việc ghép ảnh yêu cầu lựa chọn vị trí chụp ảnh phù hợp sao cho ảnh ít bị biến đổi hình học nhất. Cần xác định rõ loại ảnh panorama cần ghép để lựa chọn cách chụp phù hợp.

Các phương pháp thu nhận ảnh khác nhau có thể được sử dụng để có thể thu được hình ảnh đầu vào khác nhau rồi từ đó sẽ tạo ra các loại ảnh panorama khác nhau. Có 3 cách để chụp các tấm ảnh đầu vào là:

- Máy ảnh được đặt trên chân máy và ta vừa quay máy ảnh vừa chụp để có thể thu được ảnh đầu vào.

- Máy ảnh được để trên một tấm ván trượt, hình ảnh đầu vào thu được bằng cách vừa di chuyển tấm ván và vừa chụp. Ưu điểm của phương pháp này là đảm bảo sự ổn định và chuẩn xác của ảnh đầu vào, đảm bảo không có hoặc ít sự biến đổi hình học của ảnh, các ảnh đầu vào nằm trên cùng một đường thẳng.
- Người chụp trực tiếp giữ máy ảnh trên tay và chụp ảnh bằng cách xoay hoặc đi bộ theo một hướng vuông góc với hướng chụp của máy ảnh. Nhược điểm của cách chụp này là có thể ảnh đầu vào bị biến dạng do tác động từ người chụp như bị rung, ảnh bị nghiêng, và không nằm trên một đường thẳng.

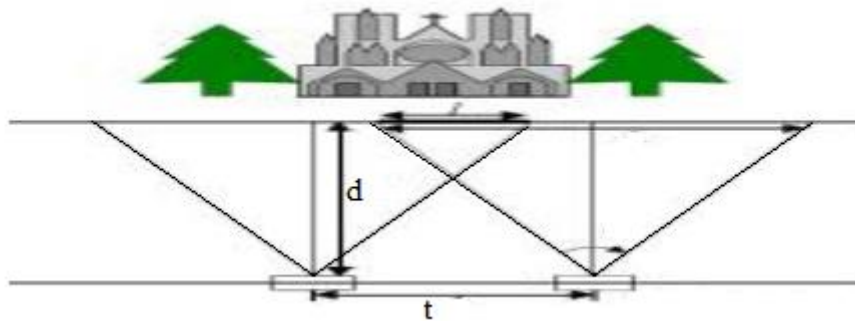
Cả ba cách chụp trên đều cần phải đảm bảo được việc ảnh sau phải chứa khoảng 15% hoặc lớn hơn nội dung của ảnh trước để đảm bảo việc xác định vị trí cần nối của hai ảnh và cần cố gắng hạn chế nhất có thể việc hình ảnh bị biến đổi dẫn đến kết quả ghép không còn chính xác nữa.

**Phương pháp thu nhận ảnh bằng camera dịch chuyển** sử dụng một máy ảnh để trên một tấm ván trượt và được di chuyển với hướng song song mặt phẳng cần chụp, máy ảnh được đặt trước đối tượng cần chụp và hình ảnh được chụp bằng cách dịch chuyển tấm ván trượt và chụp đến một giới hạn mong muốn thì thôi.





**Hình 2.4: Máy ảnh được đặt trên một giá trượt**



**Hình 2.5: Mô hình chụp ảnh có ván trượt**

Trong đó  $t$  là khoảng trượt của camera giữa hai lần chụp,  $d$  là khoảng cách giữa máy ảnh và đối tượng được chụp.

Cần đảm bảo rằng hướng trượt của máy ảnh cần phải song song với mặt phẳng chứa đối tượng cần chụp, nếu không sẽ dẫn đến việc kích thước của đối tượng sẽ bị biến đổi giữa hai ảnh.

Tuy nhiên nhược điểm của phương pháp chụp ảnh này là ảnh sau khi được ghép sẽ không cho người xem cảm giác chân thật.

**Phương pháp thu nhận ảnh bằng cách sử dụng máy ảnh cầm tay** tương đối dễ thực hiện. Người dùng chỉ cần giữ máy ảnh và vừa chụp vừa

xoay hoặc di chuyển vuông góc với hướng chụp. Tuy nhiên hình ảnh thu được thì lại khó ghép hơn do tác động như ảnh bị nghiêng, rung ...

Trong trường hợp người dùng chụp ảnh bằng cách quay người thì lúc đó người chụp đóng vai trò như là một chân máy ảnh, nhưng vẫn sẽ có sự sai lệch do tác động không mong muốn.

Trong trường hợp người dùng chụp ảnh bằng cách di chuyển song song với mặt phẳng chứa đối tượng cần chụp thì lúc đó người chụp lại đóng vai trò như là một tấm trượt. Tuy nhiên khi đó sẽ khó đảm bảo được khoảng cách ổn định từ máy ảnh đến mặt phẳng chứa đối tượng được chụp.

### **2.1.3.2 Biến đổi ảnh**

Là quá trình thay đổi về mặt hình học của ảnh để phù hợp với ảnh liền kề trước đó để tạo thành một ảnh panorama. Các ảnh cần được sắp xếp theo đúng thứ tự chụp trước sau để có thể đảm bảo được độ chính xác cao nhất có thể. Biến đổi ảnh là quá trình quan trọng nhất trong kỹ thuật ghép ảnh panorama. Quá trình bao gồm ba bước:

#### **Bước 1: Biến đổi ảnh về một không gian chiếu xác định**

Có nhiều không gian chiếu khác nhau như mặt cầu, hình trụ. Chiếu hình ảnh lên một “tấm màn chiếu” giúp biểu diễn được tấm ảnh ghép một cách chân thực hơn. Cũng có thể hiểu nó như một tấm bảng mà ta sẽ ghép các tấm ảnh lại thành ảnh panorama bằng cách dán các tấm ảnh chụp được này lên tấm bảng đó và sử dụng các phép biến đổi ảnh để thay đổi ảnh sao cho ảnh này phù hợp hơn với ảnh liền kề trước đó.

#### **Bước 2: Căn chỉnh ảnh**

Trong kỹ thuật ghép ảnh panorama thì căn chỉnh ảnh là một trong những phần quan trọng nhất. Để có thể căn chỉnh những tấm ảnh thì cần xác



định được các điểm giống nhau giữa hai bức ảnh và làm cho khoảng cách giữa hai điểm đó là ngắn nhất, thậm chí là bằng không (trường hợp trùng nhau). Để có thể làm được việc đó ta cần phải trải qua nhiều bước khác nhau.

### **Xác định điểm đặc trưng của ảnh**

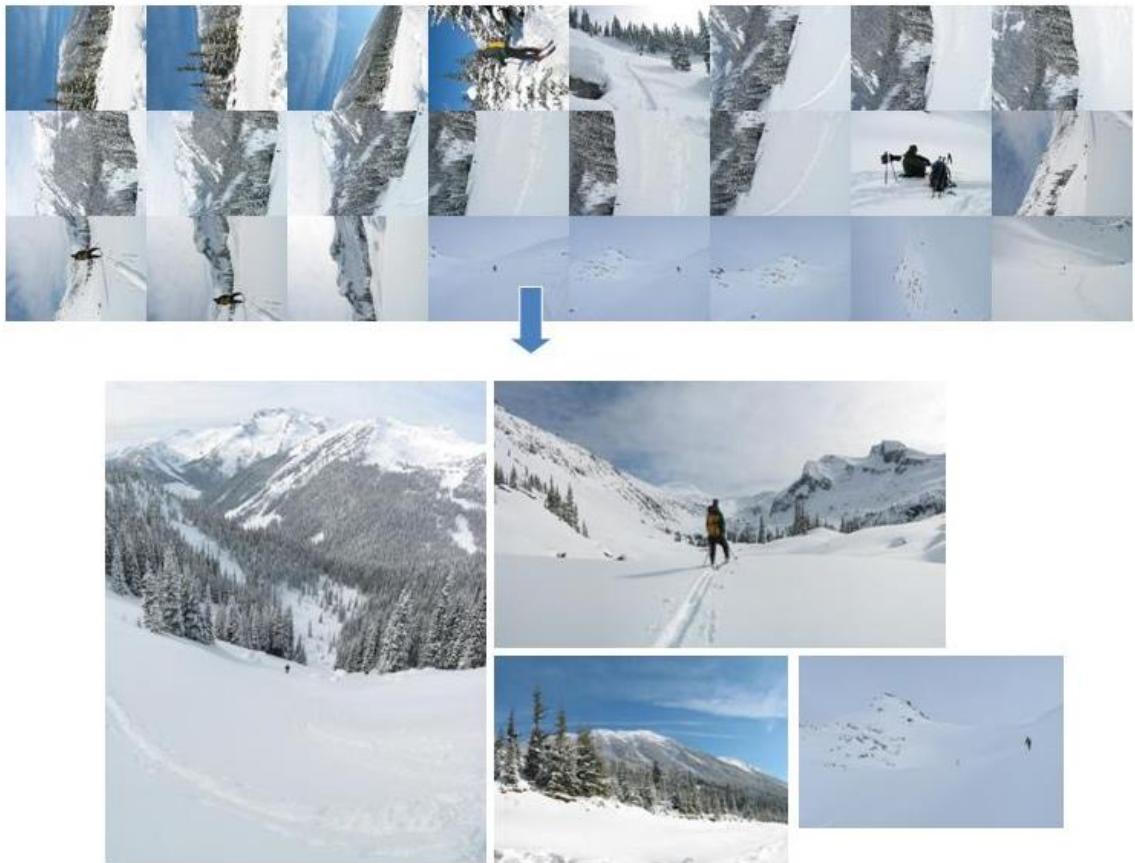
Điểm đặc trưng (hay còn gọi là key point) có thể coi là phần thông tin quan trọng được trích xuất từ ảnh và nổi bật, rõ ràng nhất của ảnh. Các điểm đặc trưng sẽ không bị thay đổi do sự biến dạng của ảnh như bị xoay hoặc do tịnh tiến ... Số lượng các điểm đặc trưng phải đủ lớn để có thể tính được một mô hình biến đổi giữa hai ảnh.

Tùy thuộc vào loại đặc trưng mà sử dụng phương pháp xác định phù hợp. Các điểm đặc trưng có thể xác định dựa vào các kỹ thuật như thuật toán phát hiện góc Harris, phát hiện cạnh...

- **Điểm đặc trưng bất biến SIFT:** Là thuật toán để mô tả các đặc trưng cục bộ dựa trên các điểm quan tâm và bất biến đối với việc co giãn hoặc xoay ảnh, không bị ảnh hưởng bởi độ sáng và nhiễu trong ảnh.
- **Tính toán ma trận đồng nhất:** Ma trận đồng nhất giữa hai ảnh có thể được tính bằng thuật toán RANSAC. Ma trận đồng nhất là ma trận thể hiện sự biến đổi của ảnh này so với ảnh kia.

### **Nhận diện ảnh panorama (Recognizing panorama)**

Trong trường hợp ảnh đầu vào không theo thứ tự hoặc là ảnh thành phần thuộc nhiều ảnh panorama khác nhau thì việc xác định từng cặp ảnh để ghép với nhau là một việc quan trọng. Kết quả trả về có thể là nhiều ảnh panorama khác nhau như **hình 2.7** Brown và Lowe đã gọi đây là kỹ thuật nhận diện ảnh panorama (recognizing panorama).



**Hình 2.6: Ví dụ cho recognize panorama**

### **Bước 3: Chiếu ảnh**

Hướng giải quyết là lựa chọn một ảnh làm tâm và biến đổi các ảnh khác theo ảnh đó. Có thể chiếu ảnh ghép lên một mặt phẳng, khi đó sẽ nhận được ảnh panorama phẳng. Cách khác là có thể sử dụng phép chiếu hình trụ (Szeliski 1994, Chen 1995) hoặc phép chiếu mặt cầu (Szeliski và Shum 1997).

#### **2.1.3.3 Trộn ảnh**

Sau khi ghép ảnh, kết quả thu được là một ảnh panorama. Tuy nhiên do tác động bên ngoài như ánh sáng, độ phơi sáng nên khi chụp ảnh đầu vào sẽ dẫn đến sự chênh lệch độ sáng màu giữa hai khu vực ảnh giống nhau giữa hai ảnh nên khi ghép vào sẽ bị nhìn rõ phần bị ghép với nhau như **hình 2.10**. Vậy cần phải làm cân bằng độ sáng của phần được ghép của hai bức ảnh để giảm đi độ rõ của vùng được nối như **hình 2.11**



**Hình 2.7: Ảnh panorama chưa được trộn màu**



**Hình 2.8: Ảnh panorama sau khi được trộn màu**

#### 2.1.3.4 Cắt ảnh

Cắt ảnh là kỹ thuật dùng để loại bỏ những đối tượng hoặc vùng ảnh bị thừa ra ngoài bức ảnh kết quả.



**Hình 2.9: Ảnh panorama kết quả**



**Hình 2.10: Ảnh panorama sau khi được cắt**

## 2.1.4 Các kỹ thuật ghép ảnh Panorama

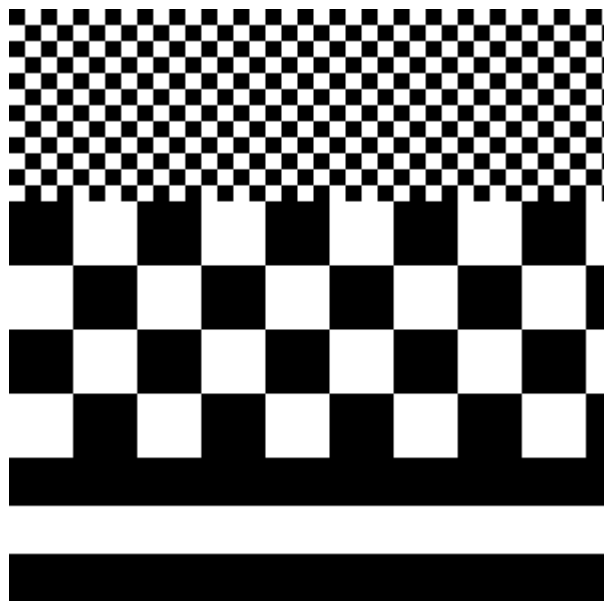
### 2.1.4.1 Ghép ảnh Panorama dựa trên kỹ thuật khớp biểu đồ màu sắc

Biến đổi hình ảnh sao cho biểu đồ màu sắc của hình ảnh đó khớp với một biểu đồ được chỉ định.

Cho hai ảnh, ảnh tham chiếu và ảnh đích. Ta tính biểu đồ histogram cho hai ảnh, ảnh tham chiếu là  $F_1()$  ảnh đích là  $F_2()$ . Sau đó với mỗi giá trị mức xám  $G_1$  trong khoảng từ 0 – 255, ta tìm giá trị mức xám tương đương là  $G_2$  sao cho  $F_1(G_1) = F_2(G_2)$  từ đó ta thu được kết quả là  $M(G_1) = G_2$ . Cuối cùng là áp dụng hàm  $M()$  cho mỗi điểm ảnh của ảnh tham chiếu.

### 2.1.4.2 Ghép ảnh dựa trên kỹ thuật phân tích kết cấu

Kết cấu hình ảnh là một tập hợp các chỉ số được tính trong xử lý ảnh được thiết kế để xác định số lượng kết cấu nhận thức của một hình ảnh. Hình ảnh kết cấu cung cấp thông tin về sự sắp xếp không gian của màu sắc hoặc cường độ trong một hình ảnh hoặc vùng lựa chọn của một hình ảnh.



**Hình 2.11: Ví dụ về kết cấu nhân tạo**





**Hình 2.12: Ví dụ về kết cấu tự nhiên**

#### **2.1.4.3 Phân tích hình khối**

Là việc sử dụng máy tính để phát hiện đối tượng có hình dạng tương tự như trong một cơ sở dữ liệu hoặc các bộ phận ăn khớp nhau. Để máy tính có thể phân tích và xử lý dạng hình học, các đối tượng phải được thể hiện dưới dạng số. Phân tích hình khối được ứng dụng trong nhiều lĩnh vực như: khảo cổ học để tìm vật tương tự hoặc các bộ phận bị thiếu, kiến trúc để xác định các đối tượng mà không gian phù hợp với một không gian cụ thể, hình ảnh y tế để hiểu những thay đổi hình dạng liên quan đến bệnh tật hoặc hỗ trợ lập kế hoạch phẫu thuật, môi trường ảo hoặc trên mô hình 3D để xác định các đối tượng vì mục đích bản quyền, các ứng dụng về bảo mật như nhận dạng khuôn mặt, ngành công nghiệp giải trí (phim ảnh, trò chơi) để xây dựng và xử lý mô hình học hoặc hình ảnh động, thiết kế hỗ trợ máy tính và máy tính hỗ trợ sản xuất để xử lý và so sánh thiết kế của các bộ phận cơ khí hoặc các đối tượng thiết kế.

#### **2.1.4.4 Ghép ảnh Panorama dựa trên kỹ thuật nắn chỉnh hình học**

Xác định độ biến dạng của ảnh thứ nhất so với ảnh thứ hai, ảnh này có thể bị tịnh tiến, co giãn theo một tỷ lệ nào đó. Công việc cần làm là biến đổi và hiệu chỉnh sao cho độ biến dạng này về mức tối thiểu.

#### **2.1.4.5 Ghép ảnh Panorama dựa trên kỹ thuật đối sánh đặc trưng của ảnh**

Thuật toán sử dụng các cặp điểm tương đồng là kết quả của thuật toán đối sánh các đặc trưng của cả hai ảnh, từ đó xây dựng nên ma trận tương đồng để có thể “chiều” bức ảnh lên một mặt phẳng trong không gian.

Phương pháp này sẽ được em trình bày chi tiết hơn ở **mục 2.2**

## **2.2 Ghép ảnh Panorama dựa trên đặc trưng bất biến của ảnh**

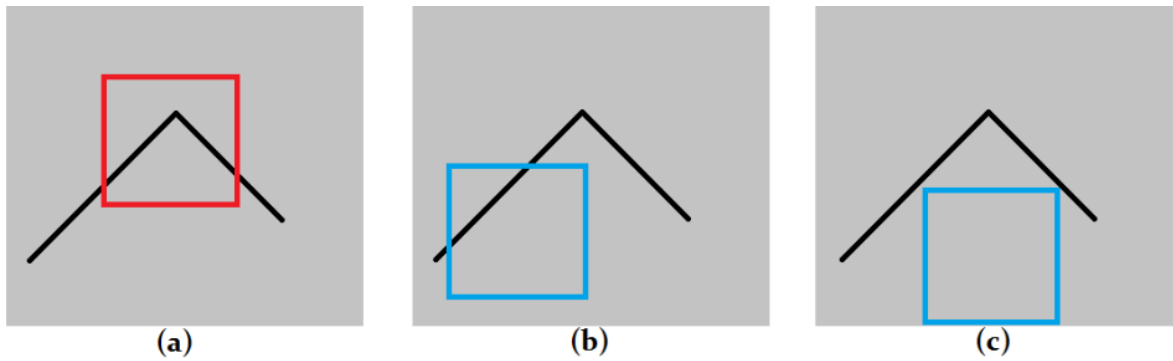
### **2.2.1 Trích chọn đặc trưng bất biến của ảnh**

Một trong những phương pháp tìm kiếm nổi bật cơ bản nhất chính là phương pháp tìm kiếm các điểm nổi của đường biên, các điểm nằm trên đường cong mà có độ cong tối đa hay còn gọi là các điểm góc.

Ban đầu các điểm này được phát hiện nhờ vào “độ nhọn” (sharp) của đường biên: biên của đối tượng được lưu dưới dạng mã xích, góc được phát hiện thông qua việc tìm kiếm những vị trí trên biên bị uốn một cách đáng kể. kỹ thuật phát hiện góc này rất phức tạp và triển khai trên nhiều bước.

Thuật toán Harris sử dụng một cửa sổ có thể trượt theo bất kỳ hướng nào bằng cách sử dụng hàm Gaussian và phép khai triển Taylor

Về mặt ý tưởng, thuật toán Harris sẽ tìm kiếm sự thay đổi lớn về cường độ xám theo các hướng khác nhau bằng cách dùng một cửa sổ nhỏ để làm nhiệm vụ rà soát và phát hiện những điểm định nghĩa là góc.



**Hình 2.13: Cửa sổ trượt phát hiện góc Harris**

Trong Hình 2.5(a) cửa sổ trượt nằm trong vùng hình ảnh có chứa góc, khi di chuyển theo bất kỳ hướng nào đều có sự thay đổi về cường độ xám.

Trong Hình 2.5(b) cửa sổ trượt nằm trên vùng hình ảnh có chứa cạnh, khi di chuyển cửa sổ trượt theo hai hướng của cạnh sẽ không có sự thay đổi về cường độ xám.

Trong Hình 2.5(c) cửa sổ trượt nằm trên vùng hình ảnh ko có góc cạnh, sau khi di chuyển cửa sổ trượt sẽ ko có sự thay đổi về cường độ xám.

Dựa vào điều này ta có thể phát hiện ra điểm nào là điểm góc, điểm nào không phải.

### 2.2.1.1 Thuật toán trích chọn đặc trưng

Giả sử cho một ảnh xám ( $I$ ), với mỗi điểm  $(u, v)$  và độ dịch chuyển  $(x, y)$  ta có thể tính toán được sự thay đổi trung bình cường độ xám bằng một cửa sổ dịch chuyển từ  $(u, v)$  tới  $(u + x, v + y)$  như sau:

$$S(x, y) = \sum_u \sum_v w(u, v) (I(u + x, v + y) - I(u, v))^2$$

Trong đó:

$S(x, y)$  là tổng bình phương giá trị độ lệch hay còn gọi là sự thay đổi cường độ xám tại  $(x, y)$

$W(u, v)$  là cửa sổ tại  $(u, v)$

$I(\mathbf{u}, \mathbf{v})$  và  $I(\mathbf{u} + \mathbf{x}, \mathbf{v} + \mathbf{y})$  là giá trị cường độ xám của pixel tại các vị trí  $(\mathbf{u}, \mathbf{v})$  và  $I(\mathbf{u} + \mathbf{x}, \mathbf{v} + \mathbf{y})$

Giá trị  $I(\mathbf{u} + \mathbf{x}, \mathbf{v} + \mathbf{y})$  có thể được khai triển theo công thức Taylor như sau:

$$I(u + x, v + y) \approx I(u, v) + I_x(u, v)x + I_y(u, v)y$$

Với  $I_x, I_y$  là đạo hàm theo thành phần  $\mathbf{x}, \mathbf{y}$ .

Từ đó, (2. 2. 1) có thể được viết lại như sau:

$$S(x, y) = \sum_u \sum_v w(u, v) (I_x(u, v)x - I_y(u, v)y)^2$$

Biểu diễn dưới dạng ma trận thì  $\mathbf{S}(\mathbf{x}, \mathbf{y})$  ta có:

$$S(x, y) \approx (x, y) \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

Trong đó,  $\mathbf{A}$  là một cấu trúc như sau:

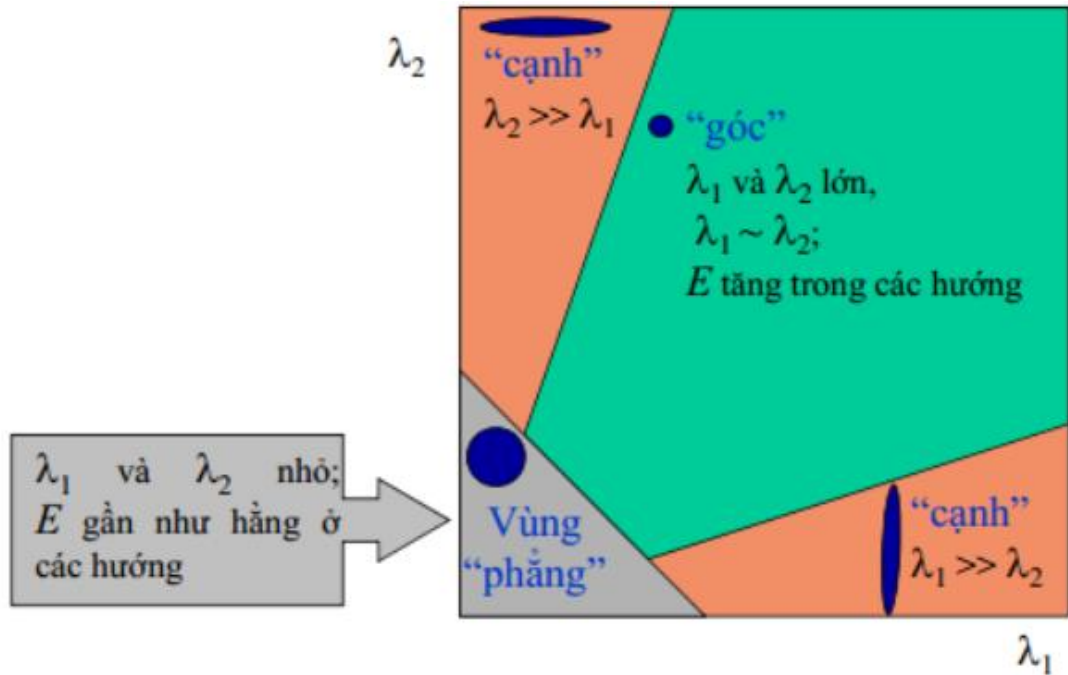
$$S(x, y) = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}$$

Gọi  $\lambda_1$  và  $\lambda_2$  là các giá trị riêng của  $\mathbf{A}$ ,  $k$  là hằng số được xác định thông qua thực nghiệm, thường có giá trị trong khoảng  $[0.04, \dots, 0.15]$ .

Khi đó, biểu thức dưới đây sẽ quyết định xem cử sở  $\mathbf{w}$  có chứa góc hay là không:

$$M_c = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 = \det(\mathbf{A}) - k(\text{trace}^2(\mathbf{A}))$$





**Hình 2.14: Minh họa các trường hợp  $\lambda_1$  và  $\lambda_2$**

Nếu cả  $\lambda_1$  và  $\lambda_2$  đều nhỏ. Có nghĩa là hàm  $S(\mathbf{x}, \mathbf{y})$  gần như không thay đổi theo bất kỳ hướng nào. Khi đó vùng ảnh nằm trong cửa sổ gần như không có sự thay đổi về cường độ. Tức là trường hợp này không tìm thấy điểm góc.

Nếu  $\lambda_1$  là lớn và  $\lambda_2$  là nhỏ hoặc ngược lại. Có nghĩa là hàm  $S(\mathbf{x}, \mathbf{y})$  có sự thay đổi nhỏ nếu cửa sổ trượt theo một hướng, và có sự thay đổi đáng kể nếu dịch chuyển theo hướng trực giao. Điều này cho thấy tồn tại một cạnh.

Nếu cả  $\lambda_1$  và  $\lambda_2$  đều lớn. Có nghĩa là hàm  $S(\mathbf{x}, \mathbf{y})$  có sự thay đổi đáng kể về cường độ xám khi dịch chuyển cửa sổ trượt theo bất kỳ hướng nào. Điều này cho thấy có tồn tại một điểm góc.

### 2.2.1.2 Xác định hướng cho các điểm nổi bật

Bằng việc gán một hướng cho mỗi điểm nổi bật dựa vào thuộc tính ảnh cục bộ, bộ mô tả điểm nổi bật có thể được biểu diễn tương đối so với hướng này và do đó đạt được tính bất biến đối với các hiện tượng quay ảnh. Độ đo của các điểm nổi bật được sử dụng để tìm ra một ảnh đã lọc Gaussian L với kích thước gần nhất sao cho mọi tính toán sẽ được thực hiện trong cùng một bất biến về độ đo.

Tính toán hướng và độ lớn cho điểm nổi bật:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1} \left( \frac{(L(x, y+1) - L(x, y-1))}{(L(x, y+1) - L(x-1, y))} \right)$$

Với:

$\mathbf{M}(x, y)$  là độ lớn của vector định hướng.

$\theta(x, y)$  là hướng của vector định hướng (biểu diễn qua góc  $\theta$ )

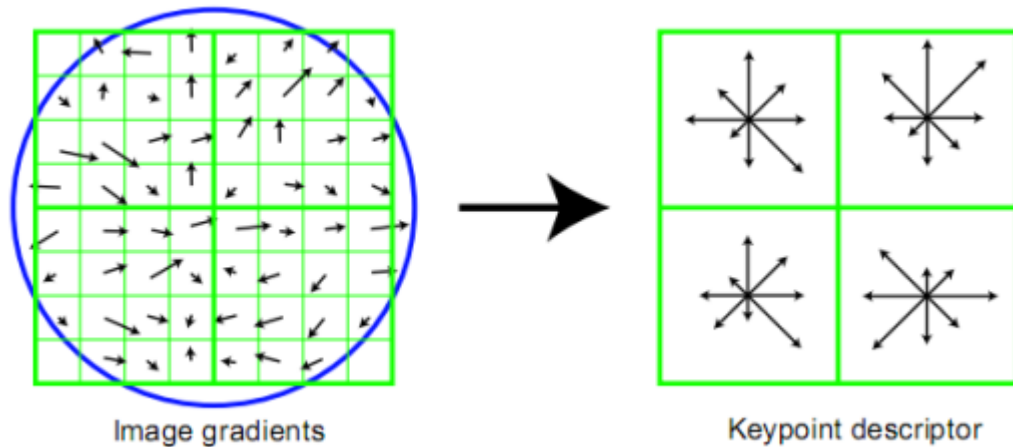
$L(x, y)$  là ảnh Gaussian ở tỷ lệ nhỏ nhất

### 2.2.1.3 Mô tả các điểm nổi bật [4]

Các phép xử lý trên đã thực hiện dò tìm và gán tọa độ, kích thước và hướng cho mỗi điểm đặc trưng. Các tham số đó yêu cầu một hệ thống tọa độ cục bộ 2D có thể lập lại được để mô tả vùng ảnh cục bộ và nhờ vậy tạo ra sự bất biến đối với các tham số đó. Bước này sẽ tính toán một bộ mô tả cho một vùng ảnh cục bộ mà có tính đặc trưng cao (bất biến với các thay đổi khác nhau về độ sáng, thu – phóng ảnh, xoay).

Một cách tiếp cận đơn giản đó là lấy mẫu mật độ ảnh cục bộ lân cận điểm đặc trưng ở một độ đo thích hợp, và đối sánh các mật độ này sử dụng độ đo tương quan chuẩn.

Cách tiếp cận tốt hơn được đưa ra bởi Edelman, Intrator và Poggio (1997). Cách tiếp cận này dựa trên một mô hình thị giác sinh học, cụ thể là mô hình neuron phức tạp trong hệ thống não bộ. Các neuron sẽ tương ứng với một gradient tại một hướng và tần số không gian cụ thể, nhưng vị trí của gradient trên võng mạc được phép trượt trên một phạm vi nhỏ của khung nhìn.



**Hình 2.15: Bộ mô tả cục bộ**

Ảnh trái là mô phỏng biên độ gradient và hướng tại mỗi mẫu ảnh trong một vùng lân cận với điểm nổi bật. Các giá trị đó tập trung trong một cửa sổ Gaussian (nằm bên trong vòng tròn). Các mẫu này sau đó được gom lại thành một lược đồ hướng mô tả vắn tắt nội dung trong 4x4 vùng con như được mô tả bên phải với độ dài của mỗi hàng tương ứng với tổng biên độ gradient gần hướng đó bên trong một vùng.

Điểm hấp dẫn sau khi được xác định hướng sẽ được biểu diễn dưới dạng các vector  $4 \times 4 \times 8 = 128$  chiều (Số chiều = 8 hướng x (4x4) điểm hấp dẫn = 128 chiều) bằng cách tổng hợp các vector định hướng của các điểm trong khu vực, các vector này có đặc điểm:

- Chung gốc.
- Độ dài mỗi vector tương ứng với độ lớn gradient  $m$  của nó.

## 2.2.2 Đối sánh các đặc trưng bất biến

### 2.2.2.1 Độ đo khoảng cách và độ đo tương tự

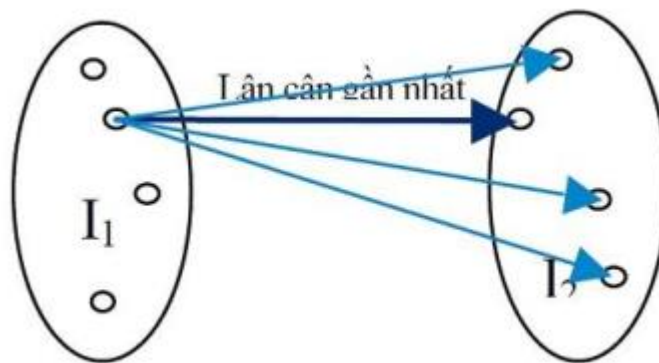
Độ đo tương tự là một trong những phương pháp tốt để máy tính phân biệt được các hình ảnh qua nội dung của chúng. Việc đối sánh theo nội dung sẽ truy vấn hình ảnh bằng phương pháp đo tương tự dựa trên các đặc trưng, việc xác định nó có thể dưới nhiều hình thức như phát hiện biên, màu sắc, vị

trí điểm ảnh ..., các phương pháp như histogram, màu sắc và phân tích sử dụng biểu đồ để xác định độ tương tự.

Do đó, độ đo có ý nghĩa quan trọng trong đối sánh ảnh dựa trên nội dung. Độ đo mang ý nghĩa quyết định kết quả đối sánh sẽ như thế nào, mức độ chính xác ra sao. Nhiều phép đo khoảng cách đã được khai thác trong việc đối sánh ảnh bao gồm: khoảng cách Euclide, khoảng cách Cousin, khoảng cách giao nhau của biểu đồ histogram, khoảng cách Minkowshi ....

### 2.2.2.2 Đối sánh đặc trưng cục bộ bất biến

Việc đối sánh sẽ được thực hiện trên các tập keypoint tìm được. Bước chính trong kỹ thuật đối sánh sẽ thực hiện tìm tập con keypoint so khớp nhau ở hai ảnh, để thực hiện việc này sẽ tìm các cặp keypoint trùng nhau lần lượt ở hai ảnh. Tập con các keypoint so khớp chính là vùng ảnh tương đồng. Việc đối sánh hai tập đặc trưng quy về bài toán tìm láng giềng gần nhất của mỗi điểm đặc trưng (hình 2. 17)



**Hình 2.16: Ví dụ về đối sánh hai tập đặc trưng**

Có 2 vấn đề cần được quan tâm:

- Tổ chức tập hợp điểm cho phép tìm kiếm láng giềng một cách hiệu quả.
- Việc đối sánh phải đạt độ chính xác nhất định.

Một phương pháp được đề xuất bởi D. Mount cho phép tìm kiếm nhanh các điểm lân cận được sử dụng, ANN được viết tắt của Approximative Nearest Neighbour. Nó cho phép tổ chức dữ liệu dưới dạng kd-tree. Cụ thể là

hai điểm trong không gian đặc trưng được coi là giống nhau nếu khoảng cách Euclide giữa hai điểm là nhỏ nhất và tỷ số giữa khoảng cách gần nhất với khoảng cách gần nhì phải nhỏ hơn một ngưỡng nào đó.

Giả sử cặp keypoint có bộ mô tả lần lượt là:

$$A = (a_1, a_2, a_3, \dots, a_{128}) \text{ và } B = (b_1, b_2, b_3, \dots, b_{128})$$

Thì khoảng cách Euclide giữa A và B được tính bằng công thức:

$$D(A, B) = \sum (a_i - b_i)^2$$

## 2.2.3 Tính toán ma trận Homography

### 2.2.3.1 Vài nét về Homography

Trong toán học, Homography là sự dịch chuyển sử dụng phép chiếu hình học, hay nói cách khác nó là sự kết hợp của cặp điểm trong phép chiếu phối cảnh. Ảnh thực trong không gian ba chiều có thể biến đổi về không gian ảnh bằng phép chiếu thông qua ma trận biến đổi Homography hay còn gọi là ma trận H. Các phép chiếu biến đổi thông qua ma trận Homography không đảm bảo về kích thước và góc của vật được chiếu nhưng lại đảm bảo về tỷ lệ.

Trong lĩnh vực thị giác máy, Homography là một ánh xạ từ mặt phẳng đối tượng đến mặt phẳng ảnh. Ma trận homography thường có liên quan đến các công việc xử lý giữa hai ảnh bất kỳ và có ứng dụng rất rộng rãi trong các công tác sửa ảnh, ghép ảnh, tính toán sự chuyển động, xoay hay dịch chuyển giữa hai ảnh.

Ta có công thức sau:

$$HX = sX'$$

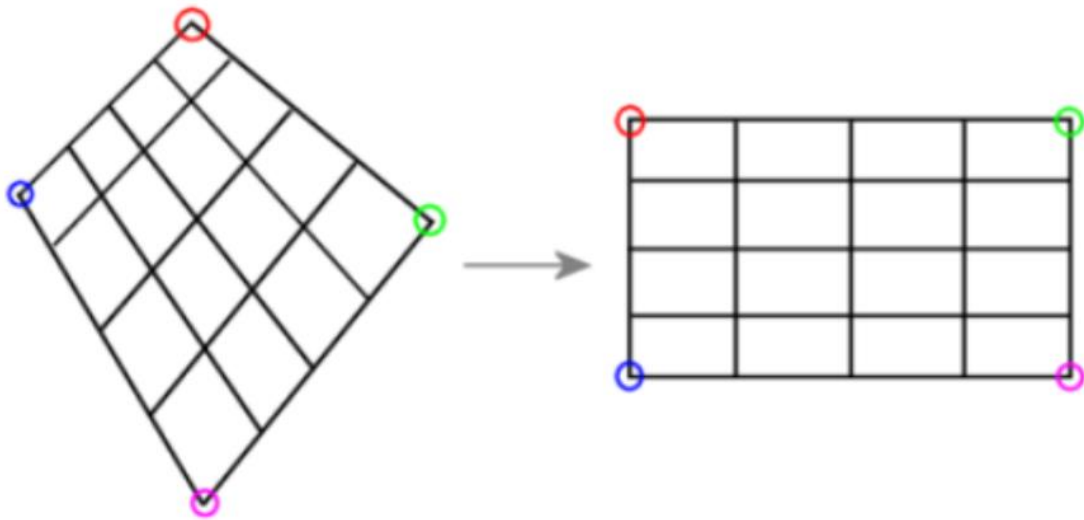
Trong đó:

S là hằng số tỉ lệ của phép chiếu và khác 0

X' là kết quả của phép ánh xạ

H là ma trận Homography, là một ma trận khả nghịch

Vì  $H$  là một ma trận khả nghịch, cho nên trong trường hợp muốn tái tạo ảnh  $X$  từ  $X'$ , chỉ cần xác định được ma trận Homography là được.



**Hình 2.17: Phép chiếu Homography**

### 2.2.3.2 Tính toán Homography

Homography là một định nghĩa toán học. Đó là sự dịch chuyển sử dụng phép chiếu hình học, hay nói cách khác nó là sự kết hợp của cặp điểm trong phép chiếu phối cảnh. Ảnh thực trong không gian ba chiều có thể biến đổi về không gian ảnh bằng phép chiếu thông qua ma trận biến đổi Homography  $H$ . Các phép chiếu biến đổi thông qua ma trận Homography tuy không đảm bảo về kích thước và góc của vật được chiếu, nhưng đảm bảo về tỷ lệ.

#### Tính ma trận Homography bằng phương pháp Direct Linear Transform

Để tính ma trận Homography từ các cặp điểm tương ứng, người ta dùng phương pháp DLT (Direct Linear Transform), gồm có 2 bước: Đầu tiên, từ các cặp điểm tương ứng, ta chuyển về dạng ma trận  $A; \mathbf{h} = \mathbf{0}$ . Sau đó, áp dụng phân rã SVD để tính ma trận  $H$ .

Giải thuật SVD hay được gọi là giải thuật phân rã giá trị đơn (Singular Value Decomposition) được Golub và Kahan công bố năm 1965 [5], đó là một kỹ thuật phân rã được sử dụng để giảm hạng (hay số chiều) của ma trận.

SVD cho phép phân tích một ma trận phức tạp thành ba ma trận thành phần. Mục đích nhằm đưa ra việc giải quyết bài toán liên quan đến ma trận lớn, phức tạp về những bài toán nhỏ hơn.

$$A = USV^T$$

Trong đó:

U là ma trận trực giao cấp  $m \times n$  (m số từ chỉ mục) các vector dòng của U là các vector từ chỉ mục.

S là ma trận đường chéo cấp  $r \times r$  có giá trị suy biến  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$  với  $r = \text{rank}(A)$

V là ma trận trực giao cấp  $r \times n$  – các vector cột của V là các vector văn bản

Hạng của ma trận A là các số dương trên đường chéo của ma trận S.

Phương pháp DLT với các điểm nổi bật được tìm thấy từ thuật toán Harris:

Trong tọa độ không đồng nhất, ta có công thức:

$$c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Lần lượt chia dòng thứ nhất của công thức trên cho dòng thứ ba và dòng thứ hai cho dòng thứ ba, ta có:

$$-h_1x - h_2y - h_3 + (h_7x + h_8y + h_9)_u = 0$$

$$-h_4x - h_5y - h_6 + (h_7x + h_8y + h_9)_u = 0$$

Viết dưới dạng ma trận ta có:

$$A_i h = \begin{bmatrix} -x & -y & -1 & 0 & 0 & 0 & ux & uy & u \\ 0 & 0 & 0 & -x & -y & -1 & vx & vy & v \end{bmatrix} (h_1 h_2 \dots h_9)^T$$

Áp dụng công thức phân rã SVD cho ma trận  $[A]$  ta có:

$$A = U \sum V^T = \sum_{i=1}^9 s_i u_i v_i^T$$

Với  $s_i$  là các giá trị đơn và được sắp xếp nhỏ dần, nên  $s_9$  là giá trị nhỏ nhất. Khi đó, giá trị của  $h_i$  bằng giá trị cuối cùng của cột  $v_i$ .

Trong thực tế các ảnh đầu vào có thể có góc tọa độ ở góc trái của ảnh, cũng có thể góc tọa độ nằm ở tâm ảnh. Nếu để tình trạng như vậy sẽ ảnh hưởng đến các kết quả biến đổi về sau như khi nhân ảnh với một hệ số hay các biến đổi tương tự affin. Các ảnh cần phải chuẩn hóa bằng phép biến đổi quay và dịch chuyển.

### Thuật toán RANSAC [6]

RANSAC (RANdom SAmple Consensus) được công bố bởi Fischler và Boller vào năm 1981. Ý tưởng chính của RANSAC như sau: Từ tập dữ liệu ban đầu, ta sẽ có hai loại dữ liệu nhiễu và không nhiễu (outlier và inlier), vì thế ta phải đi tính toán để tìm ra mô hình tốt nhất cho tập dữ liệu. Việc tính toán và chọn ra mô hình tốt nhất sẽ được lặp đi lặp lại  $k$  lần, với giá trị  $k$  được chọn sao cho đủ lớn để đảm bảo xác suất  $p$  (thường rơi vào giá trị 0.99) của tập dữ liệu mẫu ngẫu nhiên không chứa dữ liệu nhiễu.

Quá trình thực hiện thuật toán RANSAC được mô tả như dưới đây:

Từ tập dữ liệu đầu vào gồm có nhiễu và không có nhiễu ta chọn từ  $n$  dữ liệu ngẫu nhiên, tối thiểu để xây dựng mô hình:

- Tiến hành xây dựng mô hình với  $n$  dữ liệu đó, sau đó ra một ngưỡng dùng để kiểm chứng mô hình.



- Gọi tập dữ liệu ban đầu trừ đi tập  $n$  dữ liệu để xây dựng mô hình tập dữ liệu kiểm chứng. Sau đó, tiến hành kiểm chứng mô hình đã xây dựng bằng tập dữ liệu kiểm chứng. Nếu kết quả thu được từ mô hình vượt quá ngưỡng, thì điểm đó là nhiễu, còn không đó sẽ là ngược lại.
- Quá trình này sẽ được lặp lại trong  $k$  lần. Với  $k$  được tính theo công thức trên. Tại mỗi vòng lặp giá trị của  $k$  sẽ được tính lại.
- Kết quả là mô hình nào có số dữ liệu không nhiễu nhiều nhất sẽ được chọn là mô hình tốt nhất.

Trong bài toán tạo ảnh Panorama, ma trận Homography được tính từ tập các cặp điểm nổi bật tương ứng của hai ảnh ban đầu đã được so sánh đối chiếu ở bước hai. Khi đó bốn cặp điểm nổi bật tương ứng không thẳng hàng, phương trình  $Ah = 0$  theo phương pháp DLT chuẩn hóa đã trình bày ở phần trên. Trong đó,  $A$  là ma trận có kích thước  $8 * 9$ . Từ đó, ta xác định được ma trận  $h$ .

Với ma trận Homography được tính từ bốn cặp điểm ngẫu nhiên, ta có  $d$  là khoảng cách đo mức độ gần nhau của các cặp điểm đã được so sánh đối chiếu. Với cặp điểm nổi bật tương đồng  $(x, x')$  và  $d(\vec{a}, \vec{b})$  là khoảng cách của hai vector, ta có công thức khoảng cách như sau:

$$d = d(\vec{x}, H\vec{x}') + d(\vec{x}', H\vec{x})$$

Thuật toán chi tiết:

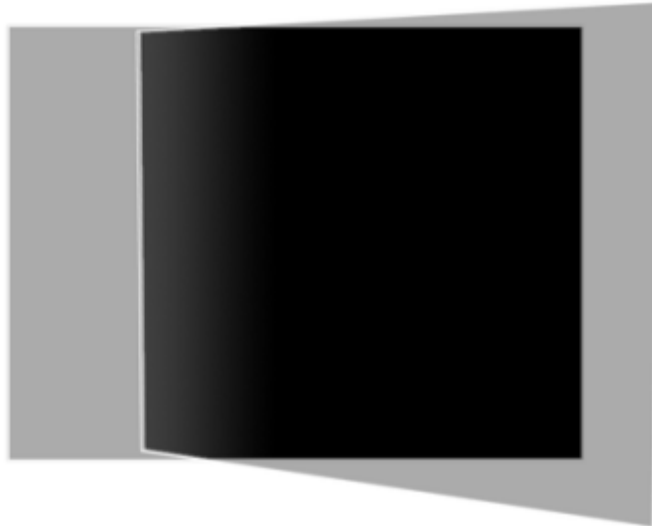
- Khởi tạo số vòng lặp  $k$ , ngưỡng *distance*,  $max_{inlier}$ , và  $p = 0$ .
- *for*( $i = 1:k$ ), thực hiện các bước sau:
  - **Bước 1:** Chọn 4 cặp điểm tương đồng ngẫu nhiên
  - **Bước 2:** Kiểm tra xem các điểm có nằm trên cùng một đường thẳng hay ko.

- **Bước 3:** Tính ma trận Homography  $H$  từ 4 điểm sử dụng phương pháp DLT chuẩn hóa.
- **Bước 4:** Tính khoảng cách  $d$  của các cặp điểm nổi bật tương đồng
- **Bước 5:** Tính số lượng  $m$  các cặp điểm không ngẫu nhiên (inlier) thỏa điều kiện:  $d_i < \text{distance}$
- **Bước 6:** Nếu  $m > \max_{\text{inlier}}$  thì  $\max_{\text{inlier}} = m$ , ma trận Homography  $H = H_{\text{curr}}$ .
- Tiếp tục tính lại ma trận  $H$  cho tất cả các cặp điểm tương đồng được coi là không nhiễu (inlier) bằng phương pháp DLT.

#### 2.2.4 Ghép ảnh dựa trên ma trận Homography

Sau khi ma trận Homography được tính toán, bước cuối cùng trong việc tạo ảnh Panorama là hòa trộn hai bức ảnh lại với nhau. Ý tưởng của bước này là sử dụng một ảnh làm trung tâm, sau đó sử dụng ma trận Homography để chiếu ảnh còn lại tới mặt phẳng ảnh trung tâm.

Như trong hình 2.18, bên trái là ảnh trung tâm được sử dụng là mặt phẳng chiếu. Bên phải là ảnh được chiếu lên mặt phẳng ảnh thứ nhất sử dụng ma trận Homography. Phần màu xám là phần riêng của mỗi bức ảnh, còn phần màu đen là phần chung của hai bức ảnh. Để thực hiện được kỹ thuật này ta dùng phép chiếu phối cảnh để chiếu các hình ảnh đầu vào ban đầu sử dụng ma trận homography làm ma trận biến đổi để chiếu lên mặt phẳng chiếu.



**Hình 2.18: Minh họa ghép nối ảnh**

Phép chiếu là phép chuyển đổi những điểm trong hệ tọa độ  $n$  chiều thành những điểm trong hệ tọa độ có số chiều ít hơn  $n$ . Một điểm  $P$  trên đối tượng được xác định trong tọa độ thế giới thực tại vị trí  $(x, y, z)$  qua phép chiếu phối cảnh  $t$  được điểm  $P'$  có tọa độ  $(x', y', z')$  nằm trên mặt phẳng chiếu.

## CHƯƠNG 3: THỰC NGHIỆM TẠO ẢNH PANORAMA

### 3.1 Môi trường cài đặt

Chương trình sử dụng ngôn ngữ lập trình matlab với mục đích triển khai các kỹ thuật sử dụng để ghép ảnh panorama.

Matlab là viết tắt từ “MATrix LABoratory”, được Clve Moler phát minh vào cuối thập niên 1970, và sau đó là chủ nhiệm khoa máy tính tại Đại học New Mexico.

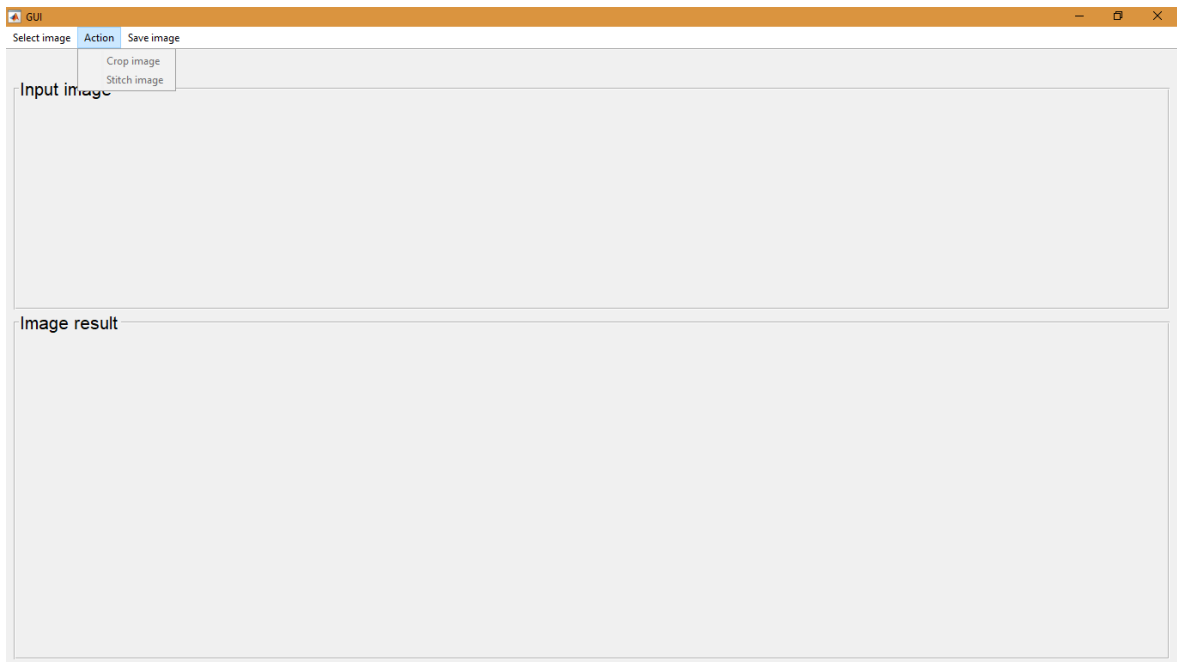
MATLAB, nguyên sơ được viết bởi ngôn ngữ Fortran, cho đến 1980 nó vẫn chỉ là một bộ phận được dùng nội bộ của Đại học Stanford.

Năm 1983, Jack Little, một người đã học ở MIT và Stanford, đã viết lại MATLAB bằng ngôn ngữ C và nó được xây dựng thêm các thư viện phục vụ cho thiết kế hệ thống điều khiển, hệ thống hộp công cụ, mô phỏng, ... Jack xây dựng MATLAB trở thành mô hình ngôn ngữ lập trình trên cơ sở ma trận.

Steve Bangert là người đã viết trình thông dịch cho MATLAB. Công việc này kéo dài gần 18 tháng. Sau này, Jack Little kết hợp với Moler và Steve Bangert quyết định đưa MATLAB thành dự án thương mại - công ty The MathWorks ra đời thời gian này - năm 1984.

Phiên bản Đầu tiên MATLAB 1. 0 ra đời năm 1984 viết bằng C do MS-DOS PC được phát hành đầu tiên tại IEEE Conference on Design and Control tại Las Vegas, Nevada. Ban đầu MATLAB được phát triển để hỗ trợ sinh viên sử dụng hai thư viện LINPACK và EISPACK dùng cho hai đại số tuyến tính mà không cần biết lập trình Fortan.

### 3.2 Giao diện chương trình



**Hình 3.1: Giao diện chính của chương trình**

Giao diện của chương trình bao gồm:

- **Select image:** Cho phép chọn tập ảnh đầu vào.
- **Crop image:** Cho phép cắt ảnh.
- **Stitch image:** Thực hiện ghép ảnh.
- **Panel input image:** nơi hiển thị ảnh đầu vào.
- **Panel image result:** nơi hiển thị ảnh kết quả bao gồm ảnh panorama.
- **Save image:** Lưu ảnh kết quả vào máy tính.

### 3.3 Chạy chương trình thực nghiệm

Trong chương trình ta sẽ tiến hành chạy thực nghiệm ghép ba ảnh đầu vào là hình 3.2, hình 3.3 và 3.4



## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến



**Hình 3.2: Ảnh đầu vào thứ nhất**



**Hình 3.3: Ảnh đầu vào thứ hai**

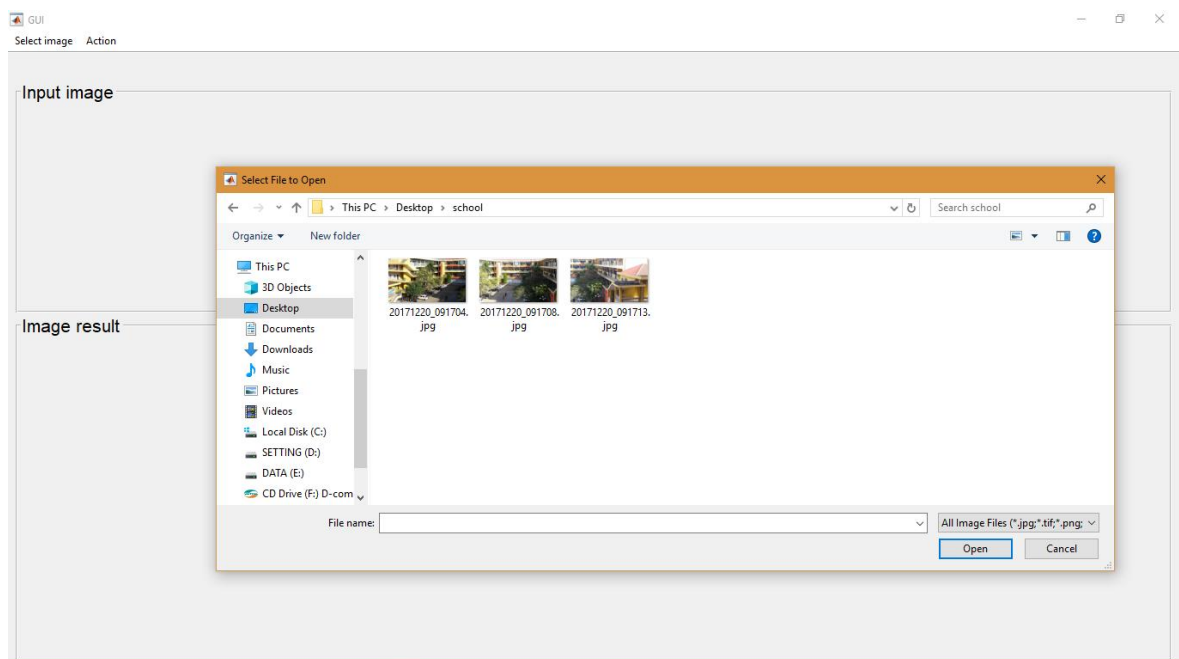


## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến



**Hình 3.4: Ảnh đầu vào thứ ba**

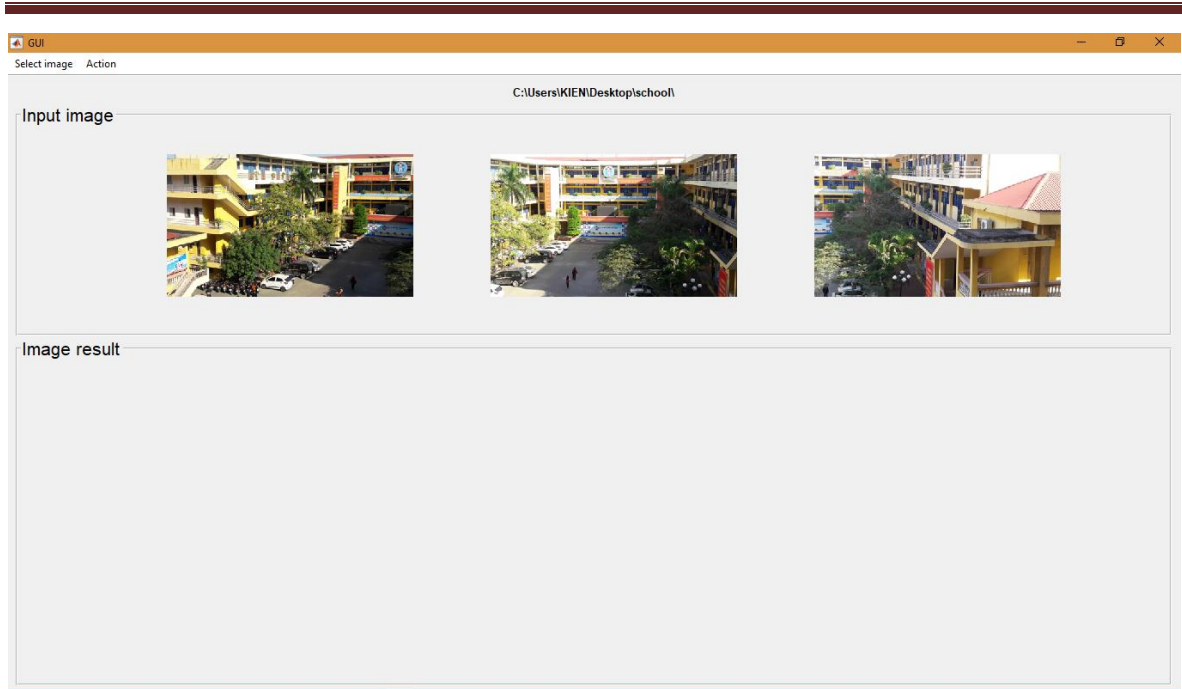
Sau khi click menu “Select image” sẽ hiện ra hộp thoại chọn ảnh



**Hình 3.5: Hộp thoại chọn ảnh để ghép**

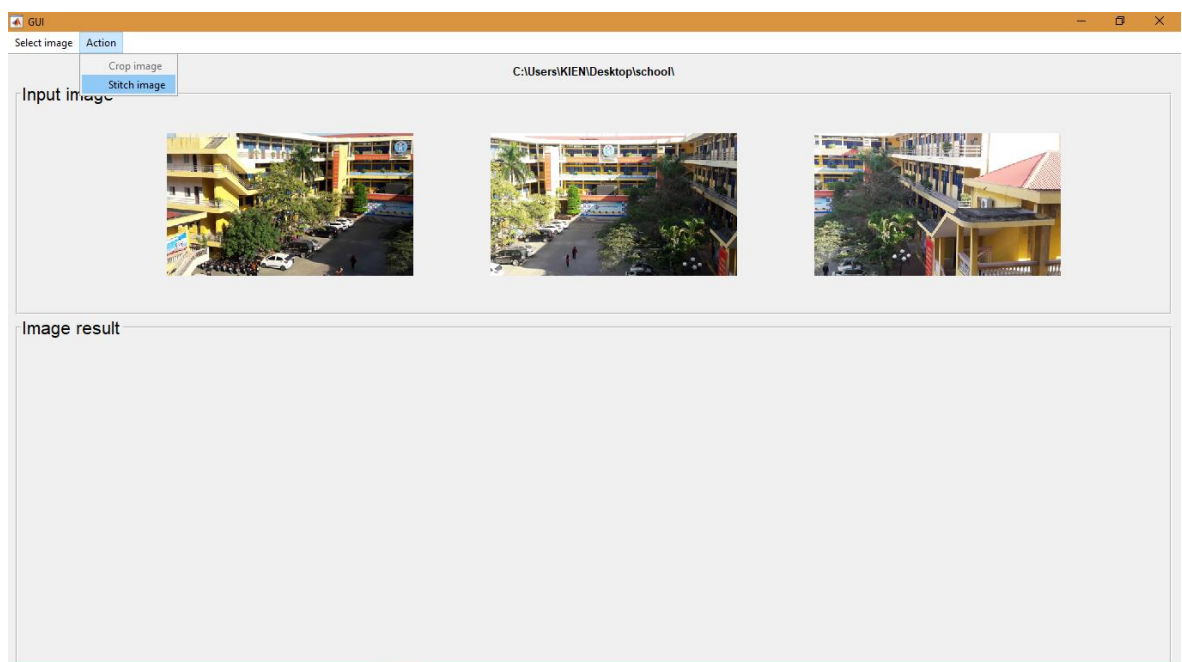
Sau khi tiến hành chọn ảnh thì ảnh được chọn sẽ hiển thị lên khu vực “input image”.

## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến



**Hình 3.6: Hình ảnh được chọn sẽ hiển thị trên giao diện**

Sau đó click menu “Action” và chọn “Stitch image” chương trình sẽ bắt đầu tiến hành ghép ảnh.



**Hình 3.7: Click “Stitch image” để tiến hành ghép ảnh**

Đầu tiên chương trình sẽ tìm kiếm các góc ở trong hai bức ảnh dựa vào thuật toán Harris. Ta có kết quả như hình 3.8, 3.9 và 3.10 với các góc được đánh dấu là các điểm ảnh màu xanh lá.





**Hình 3.8: Kết quả tìm kiếm góc cho ảnh đầu vào thứ nhất**



**Hình 3.9: Kết quả tìm kiếm góc cho ảnh đầu vào thứ hai**





**Hình 3.10: Kết quả tìm kiếm góc cho ảnh đầu vào thứ ba**

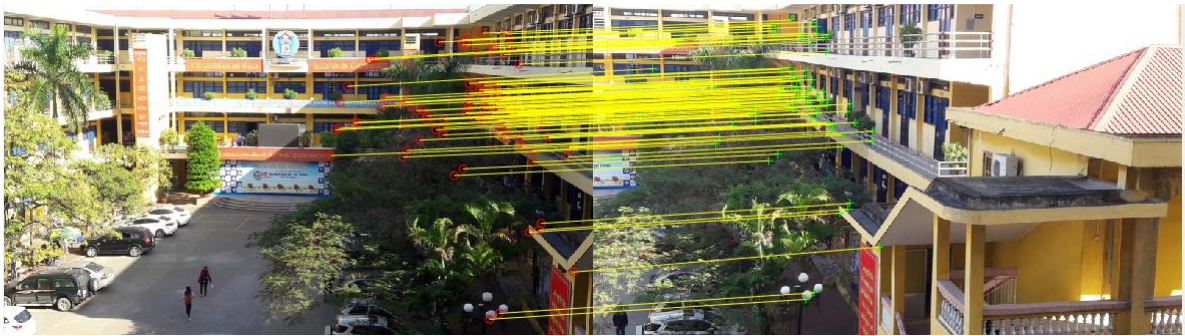
Tiếp theo, sau khi tìm được các góc trong cả hai bức ảnh. Chương trình tiến hành tính toán mô tả cho các điểm góc, kết quả trả về là ba ma trận có kích thước lần lượt là  $m \times 128$  và  $n \times 128$  và  $k \times 128$  trong đó:

- $m$ ,  $n$  và  $k$  hàng lần lượt là số góc và cũng là số key point trong ảnh đầu vào thứ nhất, thứ hai và thứ ba.
- 128 là số chiều của vector dùng để mô tả mỗi điểm góc hay còn gọi là điểm đặc trưng.

Dựa vào mô tả của tập điểm đặc trưng của ảnh thứ nhất và thứ hai. Ta tiến hành đối sánh hai ảnh dựa trên tập điểm đặc trưng. Kết quả của hàm đối sánh điểm đặc trưng như **hình 3.11** và **hình 3.12**



**Hình 3.11: Đối sánh ảnh thứ nhất và ảnh thứ hai**



**Hình 3.12** Đối sánh ảnh thứ hai và ảnh thứ ba

Sau bước đối sánh hai ảnh thì ta có tập các điểm tương đồng, bước tiếp theo là tính toán ma trận Homography dựa vào tập các cặp điểm tương đồng trên thuật toán RANSAC.

Sau khi tính được ma trận tương đồng Homography giữa hai ảnh đầu vào, chương trình sẽ chọn ảnh thứ hai làm “tâm” và biến đổi ảnh thứ nhất theo ảnh thứ hai và ảnh thứ ba theo ảnh thứ hai dựa vào ma trận Homography vừa tìm được.



**Hình 3.13:** Ảnh thứ nhất bị biến đổi theo ảnh thứ hai



**Hình 3.14:** Ảnh thứ hai làm tâm nên không bị biến đổi





**Hình 3.15: Ảnh thứ ba bị biến đổi theo ảnh thứ hai**

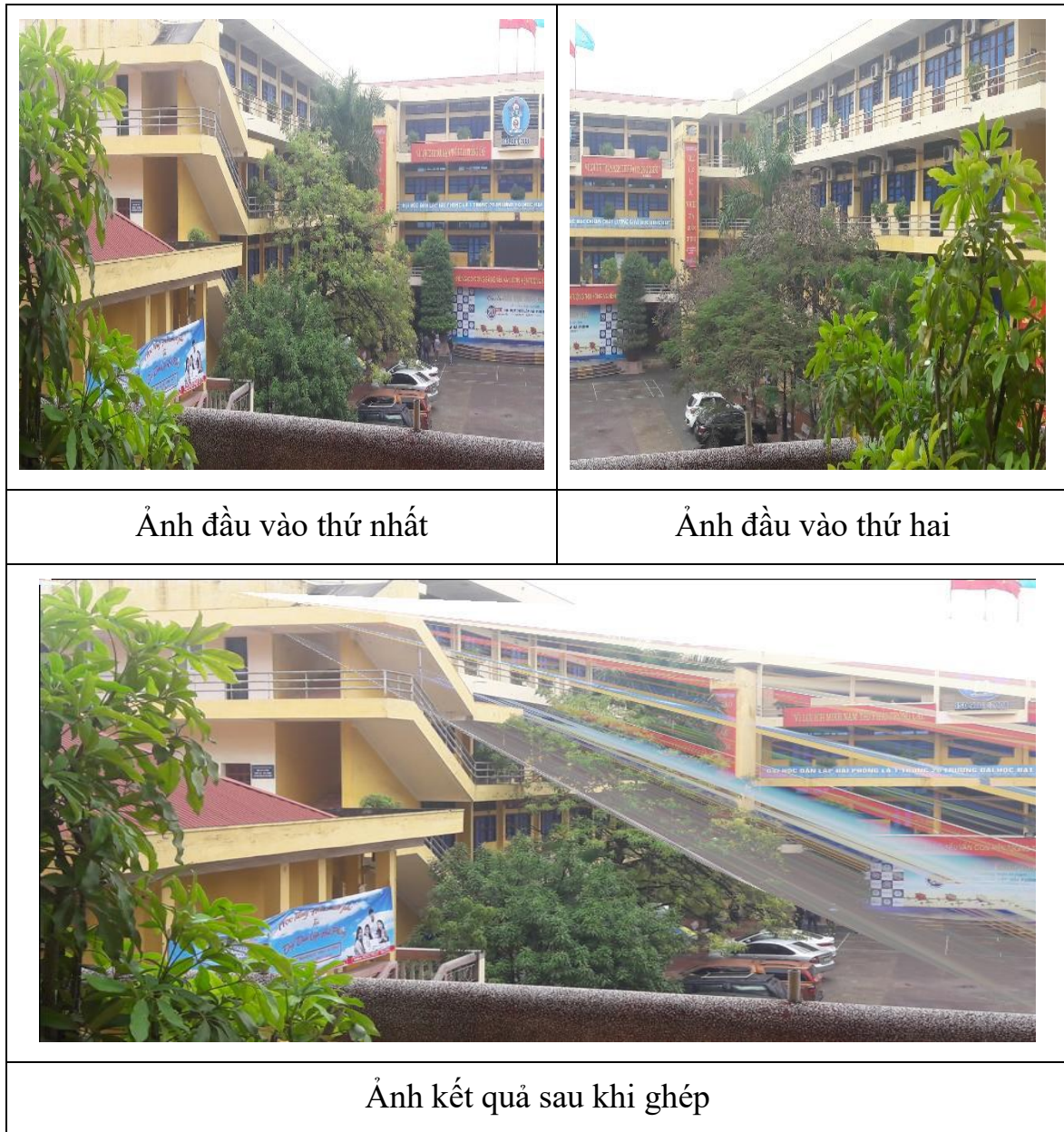
Sau khi biến đổi ảnh theo ma trận tương đồng thì bước cuối cùng của thuật toán là tiến hành chồng ba ảnh lên để tạo thành ảnh Panorama.



**Hình 3.16: Ảnh panorama kết quả**

### 3.4 Kết quả chạy thực nghiệm

Dựa vào kết quả thực nghiệm cho thấy để có thể ghép được ảnh panorama thì phải đảm bảo được nội dung ảnh thứ hai phải chứa ít nhất từ 10% đến 15 % nội dung của ảnh thứ nhất.



**Hình 3.17: Hai ảnh có tỷ lệ trùng nội dung thấp**

Nội dung của ảnh thứ hai và ảnh thứ nhất trùng lặp nhau quá ít dẫn đến việc chương trình không thể tìm ra điểm tương đồng hoặc tìm thấy quá ít dẫn đến việc tính toán ma trận homography không được chính xác.



## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

Hai ảnh có vị trí chụp khác nhau thì vẫn có thể tìm được các điểm đặc trưng tương đồng và tính toán được ma trận Homography. Tuy nhiên khi ghép ảnh một số vật thể bị che khuất ở ảnh thứ nhất nhưng không bị che khuất ở ảnh thứ hai nên khi tiến hành ghép ảnh sẽ có hiện tượng các đối tượng bị chồng lên nhau.



**Hình 3.18:** Hai ảnh đầu vào có vị trí đứng chụp khác nhau

Hai ảnh (a) và (b) được chụp ở hai vị trí khác nhau dẫn đến ảnh kết quả trả về không được chính xác.

## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

Hai ảnh có vị trí chụp khác nhau thì vẫn có thể tìm được các điểm đặc trưng tương đồng và tính toán được ma trận Homography. Tuy nhiên khi ghép ảnh một số vật thể bị che khuất ở ảnh thứ nhất nhưng không bị che khuất ở ảnh thứ hai nên khi tiến hành ghép ảnh sẽ có hiện tượng các đối tượng bị chồng lên nhau.



**Hình 3.19:** Hai ảnh đầu vào có vị trí đứng chụp khác nhau

Hai ảnh (a) và (b) được chụp ở hai vị trí khác nhau dẫn đến ảnh kết quả trả về không được chính xác.



## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

Trong trường hợp ảnh thứ hai chụp lệch chéo lên trên thì vẫn đảm bảo ảnh kết quả trả về chấp nhận được.



**Hình 3.20: Hai ảnh đầu vào có vị trí lệch nhau nhiều**

Vị trí của ảnh (b) nằm phía trên bên trái của ảnh (a) tuy nhiên vì vẫn đảm bảo được điều kiện nội dung ảnh thứ hai trùng với nội dung ảnh thứ nhất từ 10-15% nên vẫn cho ra ảnh kết quả



## Ghép ảnh panorama dựa trên đối sánh đặc trưng bất biến

Số điểm đặc trưng tương đồng giữa hai ảnh phải ít nhất là 4 để có thể đủ điều kiện để thực hiện thuật toán RANSAC, từ đó tìm ra ma trận homography. Để khắc phục được vấn đề này thì ta cần phải tăng ngưỡng sử dụng trong đối sánh ảnh để có nhiều điểm tương đồng hơn. Số điểm tương đồng giữa hai ảnh càng lớn thì ảnh kết quả trả về càng chính xác.



**Hình 3.21: Hai ảnh kết quả với hai ngưỡng đối sánh khác nhau**

Từ các trường hợp đã thử nghiệm trên, có thể kết luận rằng bước thu nhận ảnh đầu vào đóng vai trò quan trọng, quyết định đến kết quả ảnh đầu ra. Chính vì lý do đó, việc sử dụng chân máy ảnh hoặc sử dụng ván trượt khi chụp ảnh là rất cần thiết để đảm bảo chất lượng ảnh đầu vào là tốt nhất và ít bị biến đổi nhất.

## KẾT LUẬN

**Sau một thời gian tìm hiểu và nghiên cứu đề tài, đồ án đã đạt được kết quả như sau:**

- Tìm hiểu nguyên lý của kỹ thuật đối sánh ảnh.
- Tìm hiểu phương pháp tìm và trích chọn các điểm keypoint và biểu diễn đặc trưng SIFT.
- Tìm hiểu quy trình ghép ảnh panorama dựa trên đối sánh đặc trưng.
- Áp dụng thành công vào việc cài đặt chương trình thử nghiệm ghép ảnh panorama.

### **Hạn chế của đề tài:**

- Trong trích chọn đặc trưng, đồ án mới chỉ tìm hiểu thuật toán Harris, chưa có sự so sánh với các thuật toán cũng như các phương pháp đối sánh khác.
- Chưa thực hiện được kỹ thuật trộn màu và cắt ảnh sau khi ghép.
- Do phát triển trên môi trường MATLAB nên giao diện của chương trình không được đẹp, tốc độ xử lý chậm.

### **Định hướng phát triển**

- Thử nghiệm trên nhiều thuật toán trích chọn đặc trưng khác nhau để có được sự so sánh, đánh giá.
- Nghiên cứu, vận dụng các thuật toán làm mịn và cắt xén ảnh trong các bước hậu xử lý.
- Nghiên cứu và vận dụng các thuật toán tối ưu để tăng tốc độ xử lý.
- Phát triển ứng dụng tạo ảnh Panorama lên các thiết bị chụp hình kỹ thuật số, thiết bị di động chạy các hệ điều hành Android, IOS hoặc Windows Phone.

## MỘT SỐ TÀI LIỆU THAM KHẢO

- [1] Markéta Potůčková, “*Image matching and its application in photogrammetry*”.
- [2] Ms. Vrushali and S. Sakharkar, “*Image stitching techniques-an overview*”.
- [3] Prof. C. S. Gode and Ms. A. N. Ganar, “*Image retrieval by using colour, texture and shape features*”.
- [4] Herbert Bay, Andreas Es, Tinne Tuytelaars and Luc Van Gool, “*Speeded-Up Robust Features (SURF)*”.
- [5] E Garcia, “*SVD and LSI tutorial*”, MIIslita.com, 2006
- [6] Martin F., Robert B. – “*Random sample consensus*”- A paradigm for model fitting with application to image analysis and automated cartography, Communications of the ACM 24 (1981) 381–395.